

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:53:10 ; Search time 165 Seconds
(without alignments)
424.264 Million cell updates/sec

Title: US-09-445-576A-7
Perfect score: 960
Sequence: 1 EPTQPKKIVNAKDVNT.....KMFDKRCRLPYICQFGIV 181

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	100.0	181	2	AAW94256 Human tet
2	960	100.0	181	8	ADO24666 Tetranect
3	960	100.0	182	7	ADD27692 Human adi
4	960	100.0	186	5	ABG80910 Human ph
5	960	100.0	197	2	AAW94262 HGFXTN123
6	960	100.0	202	2	AAR60521 Human tet
7	960	100.0	202	4	AAM24034 Human EST
8	960	100.0	202	7	ADE09166 Novel pro
9	960	100.0	202	7	ADE34332 Tetranect
10	960	100.0	202	8	ADN04116 Antipeori
11	960	100.0	228	2	AAW94261
12	957	99.7	202	6	ABR48483 Human Pla
13	955	99.5	180	7	ADD27211 Human adi
14	936	97.5	186	5	ABG80912 Human ph
15	936	97.5	188	5	ABG80908 Human FX-
16	932.5	97.1	182	1	AAP60098 Sequence
17	914	95.2	173	7	ADD27481 Human adi
18	889	92.6	181	8	ADO24678 Tetranect
19	886	92.3	167	7	ADD27117 Human adi
20	883	92.0	181	8	ADO24677 Tetranect
21	880	91.7	181	8	ADO24676 Tetranect
22	878	91.5	173	7	ADD26862 Human adi
23	875	91.1	167	7	ADD27582 Human adi
24	875	91.1	180	2	AAW94264 HGFXTN23
25	835	87.0	157	7	ADD27429 Human adi

26	777.5	81.0	169	7	ADSO8161	Ade08161 Novel pro
27	777	80.9	186	5	ABG80916	Abg80916 Mouse Pmt
28	777	80.9	187	5	ABG80914	Abg80914 Mouse FX-
29	773.5	80.6	169	8	ABM83784	Abm83784 Human dia
30	747	77.8	152	2	AAW94265	Aaw94265 HGFXTN3 f
31	743	77.4	137	5	ABG80906	Abg80906 Human hnt
32	743	77.4	137	8	ADO24667	Ado24667 Tetranect
33	743	77.4	142	5	ABG80911	Abg80911 Human ph
34	723	75.3	143	5	ABG80909	Abg80909 Human FX-
35	719	74.9	142	5	ABG80913	Abg80913 Human ph
36	674	70.2	137	8	ADO24672	Ado24672 Tetranect
37	672	70.0	137	8	ADO24674	Ado24674 Tetranect
38	666	69.4	137	8	ADO24675	Ado24675 Tetranect
39	663	69.1	137	8	ADO24673	Ado24673 Tetranect
40	621	64.7	114	7	ADD27101	Add27101 Human adi
41	618	64.3	143	5	ABG80915	Abg80915 Mouse FX-
42	617	64.3	142	5	ABG80917	Abg80917 Mouse Pmt
43	448	46.7	89	5	ABG80918	Abg80918 Human ph
44	422.5	44.0	197	3	AB28526	Ab28526 Protein e
45	422.5	44.0	197	5	AAU82642	Aau82642 Human bre

ALIGNMENTS

RESULT 1
AAW94256
ID AAW94256 standard; protein; 181 AA.
XX
AC AAW94256;
XX
DT 20-MAR-2003 (revised)
DT 26-APR-1999 (first entry)
XX
XX
DE Human tetranectin amino acid sequence.
XX
KW Trimeric polypeptide; tetranectin trimerising structural element; TTSE;
KW fusion protein; ligand binding structure; toxin; enzyme; cytokine;
KW artificial antibody; pharmacokinetic; pharmacodynamic; gene therapy;
KW transfection; imaging; tumour; human; tetranectin.
XX
OS Homo sapiens.
XX
PN W09856906-A1.
XX
PD 17-DEC-1998.
XX
PF 11-JUN-1998; 98WO-DK000245.
XX
PR 11-JUN-1997; 97DK-00000685.
XX
(THOG/) THOGERSEN H C.
PA (ETZE/) ETZERODT M.
PA (HOLT/) HOLTET T L.
PA (GRAV/) GRAVERSEN N J H.
PA (KAST/) KASTRUP J S.
PA (NIEL/) NIELSEN B B.
PA (LARS/) LARSEN I K.
XX
PI Thogersen HC, Etzerodt M, Holtet TL, Graversen NJH, Kastrup JS;
PI Nielsen BB, Larsen IK;
XX
WPI; 1999-080897/07.
PT New monomer polypeptide constructs for diagnosis and therapy - comprise a
PT tetranectin trimerising structural element covalently linked to at least
PT one heterologous moiety for providing functional activity.
XX
PS Claim 10; Page 60; 110pp; English.
CC The invention relates to the design of trimeric polypeptides using
CC polypeptide structural elements derived from the tetranectin protein
CC family. The trimeric polypeptides constructed as a monomer polypeptide

CC construct comprise at least one tetranectin trimerizing structural
 CC element (TTSE) which is covalently linked to at least one heterologous
 CC moiety, the TTSE being capable of forming a stable complex with 2 other
 CC TTSEs, with the proviso that the heterologous moiety is different from
 CC any of the fusion proteins CIIH6XTN123, H6FXTN123, H6FXTN12, H6FXTN23
 CC (AAW94261 to AAW94264). The TTSE can be used for the construction of
 CC conjugates with heterologous moieties such as a ligand binding structure,
 CC a toxin, a detectable label, an in situ activatable substance, an enzyme,
 CC a radioactive moiety, a cytokine, a non-proteinaceous polymer, a photo
 CC cross-linking agent, or a group facilitating conjugation of the monomer
 CC polypeptide construct to a target. They can be used as vehicles for
 CC assembling antibody fragments into oligomeric or multivalent entities for
 CC generating chimeric artificial antibodies having preselected
 CC pharmacokinetic and/or pharmacodynamic properties. The constructs can be
 CC used for targeted gene therapy involving selective delivery of the
 CC material for transfection or infection of the specific population of
 CC cells. They can also be used for delivering a substance to a cell or
 CC tissue or for delivering an imaging or toxin-conjugated antibody to a
 CC tumour. They can also be used for prevention or treating a disease or for
 CC diagnosis. The TTSE provides a stable structure which can act as a
 CC vehicle for a wide variety of conjugates. The present sequence represents
 CC a human tetranectin sequence from which the TTSE can be derived. (Updated
 CC on 20-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 181 AA;

Query Match 100.0%; Score 960; DB 2; Length 181;
 Best Local Similarity 100.0%; Pred. No. 8.9e-93;
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPPTQPKKIVNAKQDVNTKMFEEELKSRDLTLAQEVALIKEQALQTVCLKGTKVHMKC 60
 DB 1 EPPTQPKKIVNAKQDVNTKMFEEELKSRDLTLAQEVALIKEQALQTVCLKGTKVHMKC 60
 QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRSVGNNEAEIWLGLNDMAAE 120
 DB 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRSVGNNEAEIWLGLNDMAAE 120
 QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENC AVLSGAANGKWFDKRCRDQLPYICQFGI 180
 DB 121 GTWVDMTGARIAYKNWETEITAPDGGKTENC AVLSGAANGKWFDKRCRDQLPYICQFGI 180
 QY 181 V 181
 DB 181 V 181

RESULT 2
 ADO24666
 ID ADO24666 standard; protein; 181 AA.
 XX
 AC ADO24666;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Tetranectin protein.

XX d; antirheumatic; antiarthritic; antipsoriatic; anti-inflammatory;
 KW antimicrobial; cytostatic; neuroprotective; gene therapy;
 KW trimeric polypeptide; monomer; cytokine; trimerizing domain;
 KW rheumatoid arthritis; psoriasis; Crohn's disease; infection; malignancy;
 KW neurodegenerat.
 XX Unidentified.
 OS
 XX WO2004039841-A2.
 PN
 XX 13-MAY-2004.
 PD
 XX 29-OCT-2003; 2003WO-DK000735.
 PF
 XX 29-OCT-2002; 2002DK-0001634.
 PR
 XX 29-OCT-2002; 2002US-0421807P.

XX (BORE-) BOREAN PHARMA AS.
 PA (OTTO/) OTTOW H K.
 XX Holtet TL, Andersen MH;
 PI WPI; 2004-376164/35.
 DR
 XX New trimeric binding polypeptides for trimeric cytokines, useful for
 PT preparing a pharmaceutical composition for the treatment of pathologies
 PT mediated by trimeric cytokines, e.g. rheumatoid arthritis, psoriasis or
 PT Crohn's disease.
 XX
 PS Example 9; SEQ ID NO 96; 96pp; English.
 XX
 CC The invention relates to a novel trimeric polypeptide comprising three
 CC monomers, each of the monomers comprising a specific binding member
 CC capable of binding a trimeric cytokine, and each of the monomers
 CC comprising a trimerizing domain. The trimeric polypeptide is useful for
 CC preparing a pharmaceutical composition that may be used for treating a
 CC pathology mediated by a trimeric cytokine, such as rheumatoid arthritis,
 CC psoriasis, Crohn's disease, infections, malignancies or neurodegenerative
 CC diseases. This sequence corresponds to a protein used to generate the
 CC trimeric polypeptide of the invention.
 XX
 SQ Sequence 181 AA;

Query Match 100.0%; Score 960; DB 8; Length 181;
 Best Local Similarity 100.0%; Pred. No. 8.9e-93;
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPPTQPKKIVNAKQDVNTKMFEEELKSRDLTLAQEVALIKEQALQTVCLKGTKVHMKC 60
 DB 1 EPPTQPKKIVNAKQDVNTKMFEEELKSRDLTLAQEVALIKEQALQTVCLKGTKVHMKC 60
 QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRSVGNNEAEIWLGLNDMAAE 120
 DB 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRSVGNNEAEIWLGLNDMAAE 120
 QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENC AVLSGAANGKWFDKRCRDQLPYICQFGI 180
 DB 121 GTWVDMTGARIAYKNWETEITAPDGGKTENC AVLSGAANGKWFDKRCRDQLPYICQFGI 180
 QY 181 V 181
 DB 181 V 181

RESULT 3
 ADD27692
 ID ADD27692 standard; protein; 182 AA.
 XX
 AC ADD27692;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Human adipocyte Selected Interacting Domain (SID) prey protein #1149.

XX Adipocyte; protein-protein interaction; protein complex;
 KW bait-prey complex; Selected Interacting Domain; SID; drug screening;
 KW drug discovery; metabolic disease; obesity; lipodystrophy;
 KW diabetes mellitus; type 2; non-insulin dependent; NIDDM;
 KW adipogenesis modulation; gene therapy; human.
 XX Homo sapiens.
 OS
 XX WO2002090544-A2.
 PN
 XX 14-NOV-2002.
 PD
 XX 03-MAY-2002; 2002WO-EP006333.
 PF
 XX 04-MAY-2001; 2001US-0288885P.

```
XX PA (HYBR-) HYBRIGENICS.
XX PA (LYNX-) LYNX THERAPEUTICS INC.
XX PI Legrain P, Whiteside S, Mao J, Khrebtukova I, Luo S;
XX DR WPI; 2003-111975/10.
XX DR N-PSDB; ADD27691.
XX PT New complex of bait and prey between two polypeptides or polynucleotides
XX PT encoding the two polypeptides of adipocytes, useful for selecting a
XX PT modulating compound that inhibits or activates protein-protein
XX PT interactions.
XX PS Claim 6; SEQ ID NO 1149; 232pp; English.
XX CC The invention relates to a bait-prey complex between two adipocyte
XX CC polypeptides, or between two polynucleotides encoding adipocyte
XX CC polypeptides. The invention also relates to selected interacting Domain
XX CC (SID) polypeptides which interact with selected bait polypeptides;
XX CC polynucleotides encoding SID polypeptides; vectors comprising SID
XX CC polynucleotides; recombinant host cells comprising an adipocyte
XX CC polynucleotide or a SID-encoding vector; a method of selecting for a
XX CC compound which modulates interactions between adipocyte polypeptides;
XX CC adipocyte modulator compounds identified using the method; a
XX CC pharmaceutical composition comprising an adipocyte modulator, or a SID-
XX CC encoding vector or host cell; and a protein chip comprising adipocyte
XX CC bait polypeptides. The bait-prey complexes of the invention are useful
XX CC for selecting a compound that inhibits or activates protein-protein
XX CC interactions between adipocyte polypeptides. The modulatory compounds
XX CC identified can be used in the treatment of metabolic diseases such as
XX CC obesity, lipodystrophy and type 2 diabetes mellitus, and in the
XX CC modulation of adipogenesis. The present sequence represents a
XX CC specifically claimed adipocyte SID prey polypeptide of the invention.
XX SQ Sequence 182 AA;
Query Match 100.0%; Score 960; DB 7; Length 182;
Best Local Similarity 100.0%; Pred. No. 8.9e-93;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPPTQPKKIVNAKDVNTKMFELKSRDLTDLAQEVALLKEQQALQTVCLKGTVMKMC 60
DB 2 EPPTQPKKIVNAKDVNTKMFELKSRDLTDLAQEVALLKEQQALQTVCLKGTVMKMC 61
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 62 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 121
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRKCRDQLPYICQFGI 180
DB 122 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRKCRDQLPYICQFGI 181
QY 181 V 181
DB 182 V 182
RESULT 4
ABG80910
ID ABG80910 standard; protein; 186 AA.
XX AC ABG80910;
XX DT 29-NOV-2002 (first entry)
XX DE Human PHTN protein.
XX KW Scaffold protein; C-type lectin-like domain; CTLD; alpha-helix;
XX KW beta-strand; connecting segment; 14loop region; tetranectin;
XX KW ligand-binding specificity; human.
XX OS Homo sapiens.

XX PN WO200248189-A2.
XX PD 20-JUN-2002.
XX PF 13-DEC-2001; 2001WO-DK000825.
XX PR 13-DEC-2000; 2000DK-00001872.
XX PR 28-FEB-2001; 2001US-0272098P.
XX PA (BORE-) BOREAN PHARMA AS.
XX PI Etzerodt M, Holtet TL, Graversen NUH, Thogersen HC;
XX DR WPI; 2002-643278/69.
XX DR N-PSDB; ABS66592.
XX PT Protein comprising a variant of model C-type lectin-like domains (CTLD),
XX PT in which alpha helices, beta-strands, connecting segments are conserved
XX PT to maintain CTLD scaffold structure, while the loop region is altered.
XX PS Claim 16; Fig 10; 168pp; English.
XX CC The present invention relates to a new protein with scaffold structure of a
XX CC C-type lectin-like domains (CTLD). The invention comprises a variant of a
XX CC model CTLD where alpha-helices and beta-strands and connecting segments
XX CC are conserved such that scaffold structure of C-type lectin-like domains
XX CC (CTLD) is substantially maintained, while the 14loop region is altered by
XX CC amino acid substitution, deletion, insertion or their combination. The
XX CC invention is useful for preparing a library of nucleotide sequences
XX CC encoding related proteins by randomising part or all of the nucleic acid
XX CC sequence encoding the loop region of its CTLD. The artificial CTLD
XX CC protein products are preferable to antibody derivatives as each binding
XX CC site is a single structurally autonomous protein domain. When used as
XX CC components of compositions to be used for in vivo diagnostic or
XX CC therapeutic purposes, artificial CTLD protein products constructed on the
XX CC basis of human CTLDs are virtually identical to the corresponding natural
XX CC CTLD protein already present in the body and are therefore less
XX CC immunogenic to the patient. They also have a smaller size, and thus
XX CC provide tissue penetration and distribution, as well as shorter half life
XX CC in circulation. Since murine and human tetranectin are identical in
XX CC structure, straightforward swapping of polypeptide segments defining
XX CC ligand-binding specificity between murine and human tetranectin
XX CC derivatives may be achieved. The present amino acid sequence represents a
XX CC human protein of the invention
XX SQ Sequence 186 AA;
Query Match 100.0%; Score 960; DB 5; Length 186;
Best Local Similarity 100.0%; Pred. No. 9.2e-93;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPPTQPKKIVNAKDVNTKMFELKSRDLTDLAQEVALLKEQQALQTVCLKGTVMKMC 60
DB 5 EPPTQPKKIVNAKDVNTKMFELKSRDLTDLAQEVALLKEQQALQTVCLKGTVMKMC 64
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 65 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 124
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRKCRDQLPYICQFGI 180
DB 125 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRKCRDQLPYICQFGI 184
QY 181 V 181
DB 185 V 185
RESULT 5
AAW94262
ID AAW94262 standard; protein; 197 AA.
XX
```

AAW94262;
20-MAR-2003 (revised)
26-APR-1999 (first entry)
H6FTXN123 fusion protein sequence.
Trimeric polypeptide; tetranectin trimerising structural element; TTSE;
fusion protein; ligand binding structure; toxin; enzyme; cytokine;
artificial antibody; pharmacokinetic; pharmacodynamic; gene therapy;
transfection; imaging; tumour; human; tetranectin; H6FTXN123.
Homo sapiens.
WO9856906-A1.
17-DEC-1998.
11-JUN-1998; 98WO-DK000245.
11-JUN-1997; 97DK-00000685.
(THOG/) THOGENSEN H C.
(ETZE/) ETZERODT M.
(HOLT/) HOLTET T L.
(GRAV/) GRAVERSEN N J H.
(KAST/) KASTRUP J S.
(NIEL/) NIELSEN B B.
(LARS/) LARSEN I K.
Thogersen HC, Etzerodt M, Holtet TL, Graversen NJH, Kastrup JS;
Nielsen BB, Larsen IK;
WPI; 1999-080897/07.
New monomer polypeptide constructs for diagnosis and therapy - comprise a
tetranectin trimerising structural element covalently linked to at least
one heterologous moiety for providing functional activity.
Disclosure; Fig 6; 110pp; English.
The invention relates to the design of trimeric polypeptides using
polypeptide structural elements derived from the tetranectin protein
family. The trimeric polypeptides constructed as a monomer polypeptide
construct comprise at least one tetranectin trimerising structural
element (TTSE) which is covalently linked to at least one heterologous
moiety, the TTSE being capable of forming a stable complex with 2 other
TTSEs, with the proviso that the heterologous moiety is different from
any of the fusion proteins CIIH6FTXN123, H6FTXN123, H6FTXN12, H6FTXN23
(AAW94261 to AAW94264). The TTSE can be used for the construction of
conjugates with heterologous moieties such as a ligand binding structure,
a toxin, a detectable label, an in situ activatable substance, an enzyme,
a radioactive moiety, a cytokine, a non-proteinaceous polymer, a photo
cross-linking agent, or a group facilitating conjugation of the monomer
polypeptide construct to a target. They can be used as vehicles for
assembling antibody fragments into oligomeric or multivalent entities for
generating chimeric artificial antibodies having preselected
pharmacokinetic and/or pharmacodynamic properties. The constructs can be
used for targeted gene therapy involving selective delivery of the
material for transfection or infection of the specific population of
cells. They can also be used for delivering a substance to a cell or
tissue or for delivering an imaging or toxin-conjugated antibody to a
tumour. They can also be used for prevention or treating a disease or for
diagnosis. The TTSE provides a stable structure which can act as a
vehicle for a wide variety of conjugates. The present sequence represents
a H6FTXN123 fusion protein sequence. The specification claims that the
heterologous moiety to which the TTSEs of the invention are linked to is
specifically different from the present fusion protein sequence. (Updated
on 20-MAR-2003 to correct PA field.)

Best Local Similarity 100.0%; Pred. No. 1e-92;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPPTOKPKKIVNAKKDVNTKMFEEELKSRDLDTLAQEVALLKEQQALQTVCLKGTVMHMC 60
DB 17 EPPTOKPKKIVNAKKDVNTKMFEEELKSRDLDTLAQEVALLKEQQALQTVCLKGTVMHMC 76
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 77 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 136
QY 121 GTWVDMTGARIAYKNWETEITAPDPGGKTENCANVLGAANGKWFDRCDQLPYICQFGI 180
DB 137 GTWVDMTGARIAYKNWETEITAPDPGGKTENCANVLGAANGKWFDRCDQLPYICQFGI 196
QY 181 V 181
DB 197 V 197
RESULT 6
AAR60521
ID AAR60521 standard; protein; 202 AA.
XX
AC AAR60521;
XX
DT 25-MAR-2003 (revised)
DT 22-MAR-1995 (first entry)
XX
DE Human tetranectin.
XX
KW Serine protease; Factor-Xa; recognition site; tetranectin;
KW fusion protein cleavage; protein folding; primer;
KW polymerase chain reaction; amplification.
XX
OS Homo sapiens.
XX
PN WO9418227-A2.
XX
PD 18-AUG-1994.
XX
PF 04-FEB-1994; 94WO-DK000054.
XX
PR 04-FEB-1993; 93DK-00000130.
PR 05-FEB-1993; 93DK-00000139.
PR 03-DEC-1993; 93WO-GB0002492.
XX
PA (DENZ-) DENZYME APS.
XX
PI Thogersen HC, Holtet TL, Etzerodt M;
XX
DR WPI; 1994-279681/34.
XX
PT Refolding of polypeptide molecules - using a cyclic process involving
denaturing and renaturing conditions to produce a correctly folded prod.
XX
PS Disclosure; Page 151-52; 202pp; English.
XX
CC cDNA encoding human tetranectin (full sequence given in AAR60521) was PCR
amplified using primers given in AAQ1274-75. Amplified cDNA was linked
to a sequence encoding the Factor-Xa cleavage site (given in AAR60503),
subcloned in vector pT7H6 so that it was linked to a hexahistidine-
encoding sequence and expressed in E. coli BL21. The fusion protein was
purified on an Ni2+-activated NTA-agarose column. A cyclic procedure was
used to obtain correctly folded recombinant protein. (Updated on 25-MAR-
2003 to correct FN field.)
XX
SQ Sequence 202 AA;
Query Match 100.0%; Score 960; DB 2; Length 202;
Best Local Similarity 100.0%; Pred. No. 1e-92;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIYNAKQDVNTKMFELKSRDLTLAGQVALLKEQQALQTVCLKGTQVHMKC 60
DB 22 EPPTQPKKIYNAKQDVNTKMFELKSRDLTLAGQVALLKEQQALQTVCLKGTQVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGSTLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGSTLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAPDPGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAPDPGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 7
AAM24034
ID AAM24034 standard; protein; 202 AA.
AC AAM24034;
DT 12-OCT-2001 (first entry)
XX Human EST encoded protein SEQ ID NO: 1559.

DE Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
KW gene therapy; nutrition.

OS Homo sapiens.
XX
XX WO200154477-A2.
XX
PD 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US002687.
XX 25-JAN-2000; 2000US-00491404.
PR 17-JUL-2000; 2000US-00617746.
PR 03-AUG-2000; 2000US-00631451.
PR 15-SEP-2000; 2000US-00663870.
XX
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Wehrman T;
XX
DR WPI; 2001-476164/51.
DR N-PSDB; AAH98693.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use.
XX
XX Claim 20; Page 1069; 1275pp; English.

PS The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention

XX Sequence 202 AA;
SQ
Query Match 100.0%; Score 960; DB 4; Length 202;
Best Local Similarity 100.0%; Pred. No. 1e-92; Indels 0; Gaps 0;
Matches 181; Conservative 0; Mismatches 0;

QY 1 EPPTQPKKIYNAKQDVNTKMFELKSRDLTLAGQVALLKEQQALQTVCLKGTQVHMKC 60
DB 22 EPPTQPKKIYNAKQDVNTKMFELKSRDLTLAGQVALLKEQQALQTVCLKGTQVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGSTLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGSTLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAPDPGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAPDPGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 8
ADE09166
ID ADE09166 standard; protein; 202 AA.
AC ADE09166;
DT 29-JAN-2004 (first entry)
XX
XX Novel protein-related contig polypeptide sequence #232.

DE novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder; contig.
XX
XX Unidentified.

XX WO2003054152-A2.
XX
XX 03-JUL-2003.
XX
XX 10-DEC-2002; 2002WO-US039555.

XX 10-DEC-2001; 2001US-0339739P.
PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 24-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.

XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
XX WPI; 2003-569235/53.

XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.

XX Disclosure; SEQ ID NO 2710; 1177pp; English.
XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present amino acid sequence was used in the
CC exemplification of the invention.

XX Sequence 202 AA;
SQ
Query Match 100.0%; Score 960; DB 7; Length 202;

Best Local Similarity 100.0%; Pred. No. 1e-92;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 81
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAOQDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAOQDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 9
ADE34332
ID ADE34332 standard; protein; 202 AA.
AC ADE34332;
XX
DT 29-JAN-2004 (first entry)
DE Tetranectin #SEQ ID 13.
XX
KW Osteopathic; vaccine; osteoarthritis; gene modulation.
XX
OS Homo sapiens.
XX
PN WO2003048394-A1.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-US038407.
XX
PR 03-DEC-2001; 2001US-0337417P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Carman J, Nadler SG, Bowen M, Neubauer M, Lu P;
XX
DR WPI; 2003-513771/48.
DR N-PSDB; ADE34331.
XX
PT Identifying a compound that modulates the activity of osteoarthritis-
associated gene or protein by determining whether the test compound
modulates the activity of the gene or protein expressed in the cell
contacted with the compound.
XX
PS Claim 10; Fig 11; 90pp; English.
XX
CC The invention relates to an assay for identifying a compound that
modulates the activity of a gene or protein associated with
osteoarthritis. The method of the invention comprises providing a cell
expressing a gene or protein associated with osteoarthritis, contacting
the cell with a test compound, and determining whether the test compound
modulates the activity of the gene or protein. The method is useful for
preparing a composition for treating osteoarthritis. Sequences given in
ADE34362-ADE34355 represent osteoarthritis genes and proteins of the
invention.
XX
SQ Sequence 202 AA;

Query Match 100.0%; Score 960; DB 7; Length 202;
Best Local Similarity 100.0%; Pred. No. 1e-92;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 81
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAOQDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAOQDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 201

Best Local Similarity 100.0%; Pred. No. 1e-92;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFEEKLSRLDTLAQEVALLKEQALQTVCLKGTVMKMC 81
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAOQDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAOQDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 201

DB 142 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202
RESULT 11
AAW94261
ID AAW94261 standard; protein; 228 AA.
AC AAW94261;
XX
XX 20-MAR-2003 (revised)
DT 26-APR-1999 (first entry)
XX
XX CIIH6FXTN123 fusion protein sequence.
XX
KW Trimeric polypeptide; tetranectin trimerising structural element; TTSE;
KW fusion protein; ligand binding structure; toxin; enzyme; cytokine;
KW artificial antibody; pharmacokinetic; pharmacodynamic; gene therapy;
KW transfection; imaging; tumour; human; tetranectin; CIIH6FXTN123.
XX
OS Homo sapiens.
XX
XX WO9856906-A1.
XX
PD 17-DEC-1998.
XX
XX 11-JUN-1998; 98WO-DK000245.
PF
XX 11-JUN-1997; 97DK-00000685.
XX
XX (THOG/) THOGERSEN H C.
PA (ETZE/) ETZERODT M.
PA (HOLT/) HOLTET T L.
PA (GRAV/) GRAVERSEN N J H.
PA (KAST/) KASTRUP J S.
PA (NIEL/) NIELSEN B B.
PA (LARS/) LARSEN I K.
XX
XX Thogersen HC, Etzerodt M, Holtet TL, Graversen NJH, Kastrup JS;
PI Nielsen BB, Larsen IK;
XX
XX WPI; 1999-080897/07.
DR
XX
XX New monomer polypeptide constructs for diagnosis and therapy - comprise a
PT tetranectin trimerising structural element covalently linked to at least
PT one heterologous moiety for providing functional activity.
XX
XX Disclosure; Fig 6; 110pp; English.
XX
XX The invention relates to the design of trimeric polypeptides using
CC polypeptide structural elements derived from the tetranectin protein
CC family. The trimeric polypeptides constructed as a monomer polypeptide
CC construct comprise at least one tetranectin trimerising structural
CC element (TTSE) which is covalently linked to at least one heterologous
CC moiety, the TTSE being capable of forming a stable complex with 2 other
CC TTSEs, with the proviso that the heterologous moiety is different from
CC any of the fusion proteins CIIH6FXTN123, H6FXTN123, H6FXTN12, H6FXTN23
CC (AAW94261 to AAW94264). The TTSE can be used for the construction of
CC conjugates with heterologous moieties such as a ligand binding structure,
CC a toxin, a detectable label, an in situ activatable substance, an enzyme,
CC a radioactive moiety, a cytokine, a non-proteinaceous polymer, a photo
CC cross-linking agent, or a group facilitating conjugation of the monomer
CC polypeptide construct to a target. They can be used as vehicles for
CC assembling antibody fragments into oligomeric or multivalent entities for
CC generating chimeric artificial antibodies having preselected
CC pharmacokinetic and/or pharmacodynamic properties. The constructs can be
CC used for targeted gene therapy involving selective delivery of the
CC material for transfection or infection of the specific population of
CC cells. They can also be used for delivering a substance to a cell or

CC tissue or for delivering an imaging or toxin-conjugated antibody to a
CC tumour. They can also be used for prevention or treating a disease or for
CC diagnosis. The TTSE provides a stable structure which can act as a
CC vehicle for a wide variety of conjugates. The present sequence represents
CC a CIIH6FXTN123 fusion protein sequence. The specification claims that the
CC heterologous moiety to which the TTSEs of the invention are linked to is
CC specifically different from the present fusion protein sequence. (Updated
CC on 20-MAR-2003 to correct PA field.)
XX
SQ Sequence 228 AA;
Query Match 100.0%; Score 960; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPPTQPKKIVNAKKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTQVHKMC 60
DB 48 EPPTQPKKIVNAKKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTQVHKMC 107
QY 61 FLAFTQTKTFHEASEDCISRGTLSTPTQGSNDALYELVRSVGNENAIWGLNDMAAE 120
DB 108 FLAFTQTKTFHEASEDCISRGTLSTPTQGSNDALYELVRSVGNENAIWGLNDMAAE 167
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRCDQLPYICQFGI 180
DB 168 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLGSAANGKWFDRCDQLPYICQFGI 227
QY 181 V 181
DB 228 V 228
RESULT 12
ABR48483
ID ABR48483 standard; protein; 202 AA.
XX
AC ABR48483;
XX
DT 13-JUN-2003 (first entry)
XX
XX Human Plasminogen carrier protein, PLCP.
DE
XX Human; GENSET; therapeutic; therapy.
KW
XX Homo sapiens.
OS
XX WO200294864-A2.
PN
XX 28-NOV-2002.
PD
XX
PF 06-AUG-2001; 2001WO-IB001715.
XX
XX 25-MAY-2001; 2001US-0293574P.
PR 15-JUN-2001; 2001US-0298698P.
PR 29-JUN-2001; 2001US-0302277P.
PR 13-JUL-2001; 2001US-0305456P.
XX
XX (GEST) GENSET.
PA
XX Bejanin S, Tanaka H;
PI WPI; 2003-129412/12.
XX N-PSDB; ACC51090.
DR
XX New GENSET polynucleotides and polypeptides, useful for preparing a
PT composition for treating GENSET-related disorders and as reagents in
PT assays to quantitatively determined levels of GENSET expression in
PT biological samples.
XX
XX Claim 2; Page 454; 505pp; English.
PS
XX The present invention relates to novel human GENSET coding sequences
CC (ACC51060-ACC51115) and proteins (ABR48453-ABR48508). The GENSET

CC sequences are useful for preparing a composition for treating GENSET-
 CC related disorders. They can also be used as markers for tissues in which
 CC the corresponding protein is preferentially expressed, as molecular
 CC weight markers on Southern gels, as chromosome markers or tags to
 CC identify chromosomes, and as reagents in assays to quantitatively
 CC determined levels of GENSET expression in biological samples
 XX
 SQ Sequence 202 AA;
 Query Match 99.7%; Score 957; DB 6; Length 202;
 Best Local Similarity 99.4%; Pred. No. 2.2e-92;
 Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEQQAQTVCCKGTVHMKC 60
 DB 22 EPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEQQAQTVCCKGTVHMKC 81
 QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
 DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
 QY 121 GTWDMTGARIAYKNWETETITAPDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 180
 DB 142 GTWDMTGARIAYKNWETETITAPDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 201
 QY 181 V 181
 DB 202 V 202
 RESULT 13
 ADD27211
 ID ADD27211 standard; protein; 180 AA.
 AC ADD27211;
 XX
 DT 15-JAN-2004 (first entry)
 DE Human adipocyte Selected Interacting Domain (SID) prey protein #675.
 DE Adipocyte; protein-protein interaction; protein complex;
 KW bait-prey complex; Selected Interacting Domain; SID; drug screening;
 KW drug discovery; metabolic disease; obesity; lipodystrophy;
 KW diabetes mellitus; type 2; non-insulin dependent; NIDDM;
 KW adipogenesis modulation; gene therapy; human.
 XX Homo sapiens.
 OS WO200290544-A2.
 PN 14-NOV-2002.
 PD 03-MAY-2002; 2002WO-EP006333.
 PF 04-MAY-2001; 2001US-0288885P.
 PR (HYBR-) HYBRIGENICS.
 PA (LYNX-) LYNX THERAPEUTICS INC.
 XX Legrain P, Whiteside S, Mao J, Khrebtukova I, Luo S;
 PI WPI; 2003-111975/10.
 DR N-PSDB; ADD27210.
 DR
 XX New complex of bait and prey between two polypeptides or polynucleotides
 PT encoding the two polypeptides of adipocytes, useful for selecting a
 PT modulating compound that inhibits or activates protein-protein
 PT interactions.
 XX Claim 6; SEQ ID NO 675; 232pp; English.
 PS The invention relates to a bait-prey complex between two adipocyte
 CC polypeptides, or between two polynucleotides encoding adipocyte

CC polypeptides. The invention also relates to Selected Interacting Domain
 CC (SID) polypeptides which interact with selected bait polypeptides;
 CC polynucleotides encoding SID polypeptides; vectors comprising SID
 CC polynucleotides; recombinant host cells comprising an adipocyte
 CC compound which modulates interactions between adipocyte polypeptides;
 CC compound which modulates interactions between adipocyte polypeptides;
 CC adipocyte modulator compounds identified using the method; a
 CC pharmaceutical composition comprising an adipocyte modulator, or a SID-
 CC encoding vector or host cell; and a protein chip comprising adipocyte
 CC bait polypeptides. The bait-prey complexes of the invention are useful
 CC for selecting a compound that inhibits or activates protein-protein
 CC interactions between adipocyte polypeptides. The modulatory compounds
 CC identified can be used in the treatment of metabolic diseases such as
 CC obesity, lipodystrophy and type 2 diabetes mellitus, and in the
 CC modulation of adipogenesis. The present sequence represents a
 CC specifically claimed adipocyte SID prey polypeptide of the invention.
 XX
 SQ Sequence 180 AA;
 Query Match 99.5%; Score 955; DB 7; Length 180;
 Best Local Similarity 100.0%; Pred. No. 3e-92;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 PPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEQQAQTVCCKGTVHMKCF 61
 DB 1 PPTQPKKIVNAKDVNTKMFEEELKSRDLTLAQEVALLEQQAQTVCCKGTVHMKCF 60
 QY 62 LAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 121
 DB 61 LAFTQTKTFHEASEDCISRGGLTSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
 QY 122 TWDMTGARIAYKNWETETITAPDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 181
 DB 121 TWDMTGARIAYKNWETETITAPDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 180
 RESULT 14
 ABG80912
 ID ABG80912 standard; protein; 186 AA.
 AC ABG80912;
 XX 29-NOV-2002 (first entry)
 DT Human Phtlec protein.
 DE Scaffold protein; C-type lectin-like domain; CTLD; alpha-helix;
 KW beta-strand; connecting segment; 14loop region; tetranectin;
 KW ligand-binding specificity; human.
 XX Homo sapiens.
 OS WO200248189-A2.
 PN 20-JUN-2002.
 PD 13-DEC-2001; 2001WO-DK000825.
 PF 13-DEC-2000; 2000DK-00001872.
 PR 28-FEB-2001; 2001US-0272098P.
 XX (BORE-) BOREAN PHARMA AS.
 XX Etzerodt M, Holtet TL, Gravarsen NJH, Thogersen HC;
 PI WPI; 2002-643278/69.
 DR N-PSDB; ABS66594.
 DR
 XX Protein comprising a variant of model C-type lectin-like domains (CTLD),
 PT in which alpha helices, beta-strands, connecting segments are conserved
 PT to maintain CTLD scaffold structure, while the loop region is altered.
 XX Claim 14; Fig 14; 168pp; English.

XX CC The present invention relates to a new protein with scaffold structure of
 CC C-type lectin-like domains (CTLD). The invention comprises a variant of a
 CC model CTLD where alpha-helices and beta-strands and connecting segments
 CC are conserved such that scaffold structure of C-type lectin-like domains
 CC (CTLD) is substantially maintained, while the 14loop region is altered by
 CC amino acid substitution, deletion, insertion or their combination. The
 CC invention is useful for preparing a library of nucleotide sequences
 CC encoding related proteins by randomising part or all of the nucleic acid
 CC sequence encoding the loop region of its CTLD. The artificial CTLD
 CC protein products are preferable to antibody derivatives as each binding
 CC site is a single structurally autonomous protein domain. When used as
 CC components of compositions to be used for in vivo diagnostic or
 CC therapeutic purposes, artificial CTLD protein products constructed on the
 CC basis of human CTLDs are virtually identical to the corresponding natural
 CC CTLD protein already present in the body and are therefore less
 CC immunogenic to the patient. They also have a smaller size, and thus
 CC provide tissue penetration and distribution, as well as shorter half life
 CC in circulation. Since murine and human tetranectin are identical in
 CC structure, straightforward swapping of polypeptide segments defining
 CC ligand-binding specificity between murine and human tetranectin
 CC derivatives may be achieved. The present amino acid sequence represents a
 CC human protein of the invention
 XX SQ Sequence 186 AA;

Query Match 97.5%; Score 936; DB 5; Length 186;
 Best Local Similarity 98.3%; Pred. No. 3.2e-90;
 Matches 178; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQQALQTVLKGTVHMKC 60
 DB 5 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQQALQTVLKGTVHMKV 64

QY 61 FLAFTQTKTFHEASDCISRGGLSTPOTGSENDALYELRQSVGNEAEIWLGLNDMAAE 120
 DB 65 FLAFTQTKTFHEASDCISRGGLSTPOTGSENDALYELRQSVGNEAEIWLGLNDMAAE 124

QY 121 GTWVDMTGARIAYKNWETEITAQPDGKGTENCANVLGSAANGKWFDRCDQLPYICQFGI 180
 DB 125 GTWVDMTGTRIAYKNWETEITAQPDGKGTENCANVLGSAANGKWFDRCDQLPYICQFGI 184

QY 181 V 181
 DB 185 V 185

RESULT 15
 ABG80908
 ID ABG80908 standard; protein; 188 AA.

XX AC ABG80908;

XX DT 29-NOV-2002 (first entry)

XX DE Human FX-htlec protein.

XX KW Scaffold protein; C-type lectin-like domain; CTLD; alpha-helix;
 KW beta-strand; connecting segment; 14loop region; tetranectin;
 KW ligand-binding specificity; human.

XX OS Homo sapiens.

XX PN WO200248189-A2.

XX PD 20-JUN-2002.

XX PF 13-DEC-2001; 2001WO-DK000825.

XX PR 13-DEC-2000; 2000DK-00001872.

XX PR 28-FEB-2001; 2001US-0272098P.

XX PA (BORE-) BOREAN PHARMA AS.

XX PI Etzerodt M, Holtet TL, Gravarsen NJH, Thogersen HC;
 XX WPI; 2002-643278/69.
 DR N-PSDB; ABS66587.
 XX PT Protein comprising a variant of model C-type lectin-like domains (CTLD),
 PT in which alpha helices, beta-strands, connecting segments are conserved
 PT to maintain CTLD scaffold structure, while the loop region is altered.

XX PS Claim 46; Fig 6; 168pp; English.

XX CC The present invention relates to a new protein with scaffold structure of
 CC C-type lectin-like domains (CTLD). The invention comprises a variant of a
 CC model CTLD where alpha-helices and beta-strands and connecting segments
 CC are conserved such that scaffold structure of C-type lectin-like domains
 CC (CTLD) is substantially maintained, while the 14loop region is altered by
 CC amino acid substitution, deletion, insertion or their combination. The
 CC invention is useful for preparing a library of nucleotide sequences
 CC encoding related proteins by randomising part or all of the nucleic acid
 CC sequence encoding the loop region of its CTLD. The artificial CTLD
 CC protein products are preferable to antibody derivatives as each binding
 CC site is a single structurally autonomous protein domain. When used as
 CC components of compositions to be used for in vivo diagnostic or
 CC therapeutic purposes, artificial CTLD protein products constructed on the
 CC basis of human CTLDs are virtually identical to the corresponding natural
 CC CTLD protein already present in the body and are therefore less
 CC immunogenic to the patient. They also have a smaller size, and thus
 CC provide tissue penetration and distribution, as well as shorter half life
 CC in circulation. Since murine and human tetranectin are identical in
 CC structure, straightforward swapping of polypeptide segments defining
 CC ligand-binding specificity between murine and human tetranectin
 CC derivatives may be achieved. The present amino acid sequence represents a
 CC human protein of the invention

XX SQ Sequence 188 AA;

Query Match 97.5%; Score 936; DB 5; Length 188;
 Best Local Similarity 98.3%; Pred. No. 3.2e-90;
 Matches 178; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQQALQTVLKGTVHMKC 60
 DB 8 EPPTQPKKIYNAKDVNTKMFELKSRDLTLAQEVALLKEQQALQTVLKGTVHMKV 67

QY 61 FLAFTQTKTFHEASDCISRGGLSTPOTGSENDALYELRQSVGNEAEIWLGLNDMAAE 120
 DB 68 FLAFTQTKTFHEASDCISRGGLSTPOTGSENDALYELRQSVGNEAEIWLGLNDMAAE 127

QY 121 GTWVDMTGARIAYKNWETEITAQPDGKGTENCANVLGSAANGKWFDRCDQLPYICQFGI 180
 DB 128 GTWVDMTGTRIAYKNWETEITAQPDGKGTENCANVLGSAANGKWFDRCDQLPYICQFGI 187

QY 181 V 181
 DB 188 V 188

Search completed: April 26, 2005, 11:57:37
 Job time : 167 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:54:54 ; Search time 140 Seconds
(without alignments)
430.251 Million cell updates/sec

Title: US-09-445-576A-7
Perfect score: 960
Sequence: 1 EPPTQPKKIVNAKKDVVNT.....KWFDRCRDQLPYICQFGIV 181

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues
Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgm2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgm2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgm2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgm2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgm2_6/ptodata/1/pubpaa/PCTU5_PUBCOMB.pep.*
7: /cgm2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgm2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgm2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgm2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgm2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgm2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgm2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgm2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgm2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgm2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgm2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	100.0	186	16	US-10-450-472-9
2	960	100.0	202	14	US-10-308-448-13
3	957	99.7	202	10	US-09-992-600A-62
4	957	99.7	202	10	US-09-924-340-62
5	957	99.7	202	10	US-09-992-095B-62
6	957	99.7	202	10	US-09-999-570-62
7	957	99.7	202	14	US-10-000-489-62
8	957	99.7	202	14	US-10-000-986-62
9	957	99.7	202	17	US-10-154-678-62
10	957	99.7	202	17	US-10-838-854-62
11	936	97.5	186	16	US-10-450-472-13
12	936	97.5	188	16	US-10-450-472-2
13	914	95.2	173	15	US-10-139-794-245
					Sequence 9, Appl
					Sequence 13, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 62, Appl
					Sequence 13, Appl
					Sequence 245, Appl

14	777	80.9	186	16	US-10-450-472-36	Sequence 36, Appl
15	777	80.9	187	16	US-10-450-472-29	Sequence 29, Appl
16	743	77.4	142	16	US-10-450-472-11	Sequence 11, Appl
17	723	75.3	143	16	US-10-450-472-4	Sequence 4, Appl
18	719	74.9	142	16	US-10-450-472-15	Sequence 15, Appl
19	618	64.4	143	16	US-10-450-472-31	Sequence 31, Appl
20	617	64.3	142	16	US-10-450-472-38	Sequence 38, Appl
21	448	46.7	89	16	US-10-450-472-57	Sequence 57, Appl
22	422.5	44.0	197	9	US-09-745-288-99	Sequence 99, Appl
23	422.5	44.0	197	14	US-10-177-293-55	Sequence 55, Appl
24	422.5	44.0	197	15	US-10-453-919-99	Sequence 99, Appl
25	422.5	44.0	206	9	US-09-989-722-403	Sequence 403, Appl
26	422.5	44.0	206	9	US-09-989-723-403	Sequence 403, Appl
27	422.5	44.0	206	9	US-09-989-279-403	Sequence 403, Appl
28	422.5	44.0	206	9	US-09-989-727-403	Sequence 403, Appl
29	422.5	44.0	206	9	US-09-989-731-403	Sequence 403, Appl
30	422.5	44.0	206	9	US-09-989-732-403	Sequence 403, Appl
31	422.5	44.0	206	9	US-09-991-073-403	Sequence 403, Appl
32	422.5	44.0	206	9	US-09-990-442-403	Sequence 403, Appl
33	422.5	44.0	206	9	US-09-991-163-403	Sequence 403, Appl
34	422.5	44.0	206	9	US-09-993-604-403	Sequence 403, Appl
35	422.5	44.0	206	9	US-09-990-456-403	Sequence 403, Appl
36	422.5	44.0	206	9	US-09-989-721-403	Sequence 403, Appl
37	422.5	44.0	206	9	US-09-992-598-403	Sequence 403, Appl
38	422.5	44.0	206	9	US-09-938-418-10	Sequence 10, Appl
39	422.5	44.0	206	9	US-09-989-293A-403	Sequence 403, Appl
40	422.5	44.0	206	9	US-09-989-735-403	Sequence 403, Appl
41	422.5	44.0	206	9	US-09-990-444-403	Sequence 403, Appl
42	422.5	44.0	206	9	US-09-991-181-403	Sequence 403, Appl
43	422.5	44.0	206	9	US-09-989-730-403	Sequence 403, Appl
44	422.5	44.0	206	9	US-09-990-436-403	Sequence 403, Appl
45	422.5	44.0	206	9	US-09-993-687-403	Sequence 403, Appl

ALIGNMENTS

RESULT 1
US-10-450-472-9
; Sequence 9, Application US/10450472
; Publication No. US20040132094A1
; GENERAL INFORMATION:
; APPLICANT: Borean Pharma A/S
; TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure
; FILE OF INVENTION: of C-type lectin-like domains
; FILE REFERENCE: BOR0003/NO
; CURRENT APPLICATION NUMBER: US/10/450,472
; CURRENT FILING DATE: 2003-12-08
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-472-9

Query Match	100.0%	Score	960;	DB	16;	Length	186;
Best Local Similarity	100.0%	Pred. No.	4.3e-92;				
Matches	181;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	EPPTQPKKIVNAKKDVVNTKMFELKSRDLDTLAEVALLKEQQALQTVCLKGTVMHMC	60				
Db	5	EPPTQPKKIVNAKKDVVNTKMFELKSRDLDTLAEVALLKEQQALQTVCLKGTVMHMC	64				
Qy	61	FLAFTQTKTFHEASDCISRGTTSTPTGSENDALYELVRSVGNELIWLGLNDMAE	120				
Db	65	FLAFTQTKTFHEASDCISRGTTSTPTGSENDALYELVRSVGNELIWLGLNDMAE	124				
Qy	121	GTWVDMTCARLAYKNWETEITTAQPDGGKTENCANVLSSGAANGKWFDRCRDQLPYICQFGI	180				
Db	125	GTWVDMTCARLAYKNWETEITTAQPDGGKTENCANVLSSGAANGKWFDRCRDQLPYICQFGI	184				
Qy	181	V	181				

```
Db      185 V 185
|
RESULT 2
US-10-308-448-13
; Sequence 13, Application US/10308448
; Publication No. US20030170743A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE DEVELOPMENT
; TITLE OF INVENTION: OSTEOARTHRITIS
; FILE REFERENCE: D0189
; CURRENT APPLICATION NUMBER: US/10/308,448
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/337,417
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-308-448-13

Query Match      100.0%; Score 960; DB 14; Length 202;
Best Local Similarity 100.0%; Pred. No. 4.8e-92;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 3
US-09-992-600A-62
; Sequence 62, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-992-600A-62

Query Match      99.7%; Score 957; DB 10; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 4
US-09-924-340-62
; Sequence 62, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-924-340-62

Query Match      99.7%; Score 957; DB 10; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAQPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202
```

```
QY 181 V 181
Db 202 V 202

RESULT 5
US-09-992-095B-62
; Sequence 62, Application US/09992095B
; Publication No. US20030157485A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US5.DIV
; CURRENT APPLICATION NUMBER: US/09/992,095B
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..21
US-09-992-095B-62

Query Match 99.7%; Score 957; DB 10; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTGVHMKC 60
Db 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTGVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPTQSGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db 82 FLAFTQTKTFHEASEDCISRGGTLSTPTQSGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCACVLSGAANGKWFDRCDQLPYICQFGI 180
Db 142 GTWVDMTGARIYKNWETEITAPDGGKTENCACVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
Db 202 V 202

RESULT 6
US-09-992-095B-62
; Sequence 62, Application US/09992095B
; Publication No. US20030157485A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: G-091US08DIV
; CURRENT APPLICATION NUMBER: US/09/999,570
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 09/924,340

Query Match 99.7%; Score 957; DB 10; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTGVHMKC 60
Db 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLDTLAQEVALLKEQQALQTVCLKGTGVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGTLSTPTQSGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db 82 FLAFTQTKTFHEASEDCISRGGTLSTPTQSGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCACVLSGAANGKWFDRCDQLPYICQFGI 180
Db 142 GTWVDMTGARIYKNWETEITAPDGGKTENCACVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
Db 202 V 202

RESULT 7
US-10-000-489-62
; Sequence 62, Application US/10000489
; Publication No. US20030092011A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: SIGNAL
; LOCATION: 1..21
US-10-000-489-62

Query Match          99.7%; Score 957; DB 14; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 142 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 8
US-10-000-986-62
; Sequence 62, Application US/10000986
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US9.DIV
; CURRENT APPLICATION NUMBER: US 10/000,986
; PRIOR FILING DATE: 2001-11-14
; PRIOR FILING DATE: 2001-08-06
; PRIOR FILING DATE: 2001-08-06
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-06-29
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..21
US-10-000-986-62

Query Match          99.7%; Score 957; DB 14; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 142 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 9
US-10-154-678-62
; Sequence 62, Application US/10154678
; Publication No. US20030162186A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 182.US1.REG
; CURRENT APPLICATION NUMBER: US 10/154,678
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -21..-1
US-10-154-678-62

Query Match          99.7%; Score 957; DB 14; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLKEQQALQTVCLKGTKVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 142 GTWVDMTGARIYKNWETEITAPDGGKTENCANVLSGAANGKWFDRKCRDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 10
US-10-838-854-62
; Sequence 62, Application US/10838854
; Publication No. US20050026182A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US 10/838,854
; PRIOR FILING DATE: 2004-05-03
; PRIOR APPLICATION NUMBER: US 10/000,489
; PRIOR FILING DATE: 2001-11-14
```

;; PRIOR APPLICATION NUMBER: US 09/924,340
;; PRIOR FILING DATE: 2001-08-06
;; PRIOR APPLICATION NUMBER: PCT/IB01/01715
;; PRIOR FILING DATE: 2001-08-06
;; PRIOR APPLICATION NUMBER: US 60/305,456
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: US 60/302,277
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: US 60/298,698
;; PRIOR FILING DATE: 2001-06-15
;; PRIOR APPLICATION NUMBER: US 60/293,574
;; PRIOR FILING DATE: 2001-05-25
;; NUMBER OF SEQ ID NOS: 112
;; SOFTWARE: Jpatent
;; SEQ ID NO 62
;; LENGTH: 202
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SIGNAL
;; LOCATION: 1..21
US-10-838-854-62

Query Match 99.7%; Score 957; DB 17; Length 202;
Best Local Similarity 99.4%; Pred. No. 9.9e-92;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVLTGKTVHMKC 60
DB 22 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVLTGKTVHMKC 81
QY 61 FLAFTQTKTFHEASEDCISRGGLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 82 FLAFTQTKTFHEASEDCISRGGLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
DB 142 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 201
QY 181 V 181
DB 202 V 202

RESULT 11
US-10-450-472-13
;; Sequence 13, Application US/10450472
;; Publication No. US20040132094A1
;; GENERAL INFORMATION:
;; APPLICANT: Boreau Pharma A/S
;; TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure
;; FILE REFERENCE: BOR0003/WO
;; CURRENT APPLICATION NUMBER: US/10/450,472
;; NUMBER OF SEQ ID NOS: 91
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 13
;; LENGTH: 186
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-450-472-13
Query Match 97.5%; Score 936; DB 16; Length 186;
Best Local Similarity 98.3%; Pred. No. 1.4e-89;
Matches 178; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVLTGKTVHMKC 60
DB 5 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVLTGKTVHMKV 64
QY 61 FLAFTQTKTFHEASEDCISRGGLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120

DB 65 FLAFTQTKTFHEASEDCISRGGLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 124
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
DB 125 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 184
QY 181 V 181
DB 185 V 185

RESULT 12
US-10-450-472-2
;; Sequence 2, Application US/10450472
;; Publication No. US20040132094A1
;; GENERAL INFORMATION:
;; APPLICANT: Boreau Pharma A/S
;; TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure
;; FILE REFERENCE: BOR0003/WO
;; CURRENT APPLICATION NUMBER: US/10/450,472
;; NUMBER OF SEQ ID NOS: 91
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 2
;; LENGTH: 188
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-450-472-2

Query Match 97.5%; Score 936; DB 16; Length 188;
Best Local Similarity 98.3%; Pred. No. 1.4e-89;
Matches 178; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVLTGKTVHMKC 60
DB 8 EPPTQPKKIVNAKDVVNTKMFELKSRDLTLAQEVALLEQQAQTVCVLTGKTVHMKV 67
QY 61 FLAFTQTKTFHEASEDCISRGGLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 68 FLAFTQTKTFHEASEDCISRGGLSTPQTGSENDALYEYLRQSVGNEAEIWLGLNDMAAE 127
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 180
DB 128 GTWVDMTGARIAYKNWETEITAPDGGKTENCANVLSGAANGKWFDRCDQLPYICQFGI 187
QY 181 V 181
DB 188 V 188

RESULT 13
US-10-139-794-245
;; Sequence 245, Application US/10139794
;; Publication No. US20030232421A1
;; GENERAL INFORMATION:
;; APPLICANT: HYBRIGENICS, LYNX THERAPEUTICS INC.
;; APPLICANT: Pierre Legrain, Simon Whiteside, Jen-I Mao, Irina Khrebukova, Shujun Luo
;; TITLE OF INVENTION: Protein-Protein Interactions in Adipocyte Cells (3)
;; FILE REFERENCE: B4883A
;; CURRENT APPLICATION NUMBER: US/10/139,794
;; PRIOR FILING DATE: 2002-05-06
;; PRIOR APPLICATION NUMBER: US 60/288,885
;; NUMBER OF SEQ ID NOS: 2930
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 245
;; LENGTH: 173
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; OTHER INFORMATION: Translation of SEQ ID NO:242
US-10-139-794-245

```
Query Match      95.2%; Score 914; DB 15; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.5e-87;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 KIVNAKDVNTKMFELKSRDLTLAQEVALLKEQQAALQTVCLKGTVKVMKCFLLAFTQTK 68
DB 1 KIVNAKDVNTKMFELKSRDLTLAQEVALLKEQQAALQTVCLKGTVKVMKCFLLAFTQTK 60

QY 69 TTHEASEDCISRGGLTSTPQTSSENDALYEYLRQSVGNEAEIWLGLNDMAAEGTWVDMTG 128
DB 61 TTHEASEDCISRGGLTSTPQTSSENDALYEYLRQSVGNEAEIWLGLNDMAAEGTWVDMTG 120

QY 129 ARIAYKNWETEITTAQPDGKTECNCAVLGAANGKWFDFKRCRDQLPYICQFIV 181
DB 121 ARIAYKNWETEITTAQPDGKTECNCAVLGAANGKWFDFKRCRDQLPYICQFIV 173

RESULT 14
US-10-450-472-36
; Sequence 36, Application US/10450472
; Publication No. US20040132094A1
; GENERAL INFORMATION:
; APPLICANT: Borean Pharma A/S
; TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure
; FILE REFERENCE: BOR0003/WO
; CURRENT APPLICATION NUMBER: US/10/450,472
; CURRENT FILING DATE: 2003-12-08
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-450-472-36
```

```
Query Match      80.9%; Score 777; DB 16; Length 186;
Best Local Similarity 80.1%; Pred. No. 6.1e-73;
Matches 145; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFELKSRDLTLAQEVALLKEQQAALQTVCLKGTVKVMK 60
DB 5 ESPTPKAKAANAKKDLVSSKMFELKRMVDVLAQEVALLKEQQAALQTVCLKGTVKVLKV 64

QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTSSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 65 LLAFQPKTFHEASEDCISQGGTLGTPQSELENEALFEYARHSVGNDAEIVLGLNDMAAE 124

QY 121 GTWVDMTGARIAYKNWETEITTAQPDGKTECNCAVLGAANGKWFDFKRCRDQLPYICQFGI 180
DB 125 GAWVDMTGTLAYKNWETEITTPDGGKAENCAALSGAANGKWFDFKRCRDQLPYICQFAI 184

QY 181 V 181
DB 185 V 185
```

```
RESULT 15
US-10-450-472-29
; Sequence 29, Application US/10450472
; Publication No. US20040132094A1
; GENERAL INFORMATION:
; APPLICANT: Borean Pharma A/S
; TITLE OF INVENTION: Combinatorial libraries of proteins having the scaffold structure
; FILE REFERENCE: BOR0003/WO
; CURRENT APPLICATION NUMBER: US/10/450,472
; CURRENT FILING DATE: 2003-12-08
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 187
```

```
TYPE: PRT
; ORGANISM: Mus musculus
US-10-450-472-29

Query Match      80.9%; Score 777; DB 16; Length 187;
Best Local Similarity 80.1%; Pred. No. 6.1e-73;
Matches 145; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVNTKMFELKSRDLTLAQEVALLKEQQAALQTVCLKGTVKVMK 60
DB 7 ESPTPKAKAANAKKDLVSSKMFELKRMVDVLAQEVALLKEQQAALQTVCLKGTVKVLKV 66

QY 61 FLAFTQTKTFHEASEDCISRGGLTSTPQTSSENDALYEYLRQSVGNEAEIWLGLNDMAAE 120
DB 67 LLAFQPKTFHEASEDCISQGGTLGTPQSELENEALFEYARHSVGNDAEIVLGLNDMAAE 126

QY 121 GTWVDMTGARIAYKNWETEITTAQPDGKTECNCAVLGAANGKWFDFKRCRDQLPYICQFGI 180
DB 127 GAWVDMTGTLAYKNWETEITTPDGGKAENCAALSGAANGKWFDFKRCRDQLPYICQFAI 186

QY 181 V 181
DB 187 V 187

Search completed: April 26, 2005, 12:12:13
Job time : 142 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:53:10 ; Search time 42 Seconds
(without alignments)
321.702 Million cell updates/sec

Title: US-09-445-576A-7
Perfect score: 960
Sequence: 1 BPPTQPKKIIVNAKDVNT.....KWFDKRCRDQLPYICQFQIV 181

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	960	100.0	202	1 US-08-469-486-56	Sequence 56, Appl
2	960	100.0	202	2 US-08-469-658-56	Sequence 56, Appl
3	957	99.7	202	4 US-10-000-489-62	Sequence 62, Appl
4	422.5	44.0	197	4 US-09-602-877A-99	Sequence 99, Appl
5	222.5	23.2	328	4 US-09-065-040-12	Sequence 12, Appl
6	221.5	23.1	328	4 US-09-065-040-8	Sequence 8, Appl
7	220.5	23.0	323	4 US-09-065-040-4	Sequence 4, Appl
8	220.5	23.0	323	4 US-09-949-016-6294	Sequence 6294, Ap
9	189	19.7	59	4 US-09-513-999C-4342	Sequence 4342, Ap
10	186.5	19.4	1257	1 US-08-340-428B-49	Sequence 49, Appl
11	185.5	19.3	248	3 US-09-198-603C-2	Sequence 2, Appl
12	185.5	19.3	248	4 US-09-600-932-27	Sequence 27, Appl
13	185.5	19.3	248	4 US-09-949-016-5907	Sequence 5907, Ap
14	185.5	19.3	263	4 US-09-949-016-9370	Sequence 9370, Ap
15	178.5	18.6	161	3 US-09-011-735-6	Sequence 6, Appl
16	178.5	18.6	351	3 US-09-011-735-1	Sequence 1, Appl
17	178.5	18.6	351	3 US-09-029-156-1	Sequence 1, Appl
18	177	18.4	912	5 PCT-US95-03747-2	Sequence 2, Appl
19	172.5	18.0	375	4 US-09-600-932-29	Sequence 29, Appl
20	166	17.3	287	1 US-08-365-103B-4	Sequence 4, Appl
21	166	17.3	300	1 US-08-365-103B-6	Sequence 6, Appl
22	166	17.3	327	1 US-08-365-103B-2	Sequence 2, Appl
23	165	17.2	1479	3 US-08-840-062-4	Sequence 4, Appl
24	163	17.0	1456	4 US-09-976-594-168	Sequence 168, App
25	160	16.7	1479	3 US-08-840-062-2	Sequence 2, Appl
26	152	15.8	174	2 US-08-401-530A-5	Sequence 5, Appl
27	152	15.8	174	2 US-08-709-662-5	Sequence 5, Appl

28	151	15.7	1487	3 US-08-840-062-7	Sequence 7, Appli
29	149.5	15.6	1326	4 US-09-949-016-6806	Sequence 6806, Ap
30	148.5	15.5	1339	4 US-09-949-016-10448	Sequence 10448, A
31	148.5	15.5	1478	4 US-09-949-016-8315	Sequence 8315, Ap
32	148	15.4	125	3 US-08-722-136A-7	Sequence 7, Appli
33	148	15.4	125	5 PCT-US95-04258-7	Sequence 7, Appli
34	148	15.4	166	2 US-08-729-103-4	Sequence 4, Appli
35	148	15.4	174	4 US-09-949-016-10686	Sequence 10686, A
36	148	15.4	2409	6 5180808-2	Patent No. 5180808
37	148	15.4	2409	6 5180808-2	Patent No. 5180808
38	142	14.8	108	6 5514582-16	Patent No. 5514582
39	142	14.8	108	6 5514582-16	Patent No. 5514582
40	141	14.7	455	4 US-09-866-028-50	Sequence 50, Appl
41	141	14.7	455	4 US-09-866-028-50	Sequence 50, Appl
42	140.5	14.6	187	4 US-09-535-521-17	Sequence 17, Appl
43	140.5	14.6	208	4 US-09-535-521-20	Sequence 20, Appl
44	140.5	14.6	292	4 US-09-535-521-2	Sequence 2, Appli
45	140.5	14.6	292	4 US-09-535-521-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-469-486-56
; Sequence 56, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Ezerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-486-56

Query Match 100.0%; Score 960; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 5.3e-102;

Matches	181,	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	EPPTQPKKIYNAKDVVNTKMFELKSRDLTLAQEVALLEKEQALQTVCLKGTGVHMK	60						
Db	22	EPPTQPKKIYNAKDVVNTKMFELKSRDLTLAQEVALLEKEQALQTVCLKGTGVHMK	81						
Qy	61	FLAFTQTKTTHAESDCISRGGTLLSTPTQSGENDALYEYLRQSVGNBAEITWGLNDMAAE	120						
Db	82	FLAFTQTKTTHAESDCISRGGTLLSTPTQSGENDALYEYLRQSVGNBAEITWGLNDMAAE	141						
Qy	121	GTWVDMTGARIAYKNWETEITAO PDGCKTENC AVLSGAANGKMPDKRCRDLQPYICQFI	180						
Db	142	GTWVDMTGARIAYKNWETEITAO PDGCKTENC AVLSGAANGKMPDKRCRDLQPYICQFI	201						
Qy	181	V 181.							
Db	202	V 202							

```

RESULT 2
US-08-469-658-56
; Sequence 56, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Th egersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Eterodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOLDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25

```

Query Match 100.0%; Score 960; DB 2; Length 202;
Best Local Similarity 100.0%; Pred. No. 5.3e-102;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPTQRPKKIVNAKQDVVVNTMFEELKSRLDTTLAQEVALLKEQALQTVCLKGTFVEMKC 60

Db	22	EPPTQPKKIVNAKXDVVNTKMFEEKSRSLDTLAEQVALLKEQALQTVCLKGTKVHMK	81
Qy	61	FLAFTQTKTFHEASEDCISRGGTILSTPQTSGENDALYEYLRQSVGNEAEIWLGLNDMAAE	120
Db	82	FLAFTQTKTFHEASEDCISRGGTILSTPQTSGENDALYEYLRQSVGNEAEIWLGLNDMAAE	141
Qy	121	GTWDMTGARIYKKNWETEITAPDGGKTENC AVLSCAANGKWFDKCRDQLPYICQFGI	180
Db	142	GTWDMTGARIYKKNWETEITAPDGGKTENC AVLSCAANGKWFDKCRDQLPYICQFGI	201
Qy	181	V 181	
Db	202	V 202	

RESULT 3
 US-10-000-489-62
 ; Sequence 62, Application US/10000489
 ; Patent No. 6794363
 ; GENERAL INFORMATION:
 ; APPLICANT: Benjamin, Stephane
 ; APPLICANT: Tanaka, Hiroaki
 ; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
 ; FILE REFERENCE: 91.US6.DIV
 ; CURRENT APPLICATION NUMBER: US/10/000,489
 ; PRIOR FILING DATE: 2001-11-14
 ; PRIOR APPLICATION NUMBER: US 09/924,340
 ; PRIOR FILING DATE: 2001-08-06
 ; PRIOR APPLICATION NUMBER: PCT/IB01/01715
 ; PRIOR FILING DATE: 2001-08-06
 ; PRIOR APPLICATION NUMBER: US 60/305,456
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 60/302,277
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 60/298,698
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 60/293,574
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 112
 ; SOFTWARE: Jpatent
 ; SEQ ID NO 62
 ; LENGTH: 202
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIGNAL
 ; LOCATION: 1..21
 US-10-000-489-62

Query Match	99.7%; Score 957; DB 4; Length 202;
Best Local Similarity	99.4%; Pred. No. 1.2e-101;
Matches 180; Conservative	1; Mismatches 0; Indels 0; Gaps 0
QY	1 EPPTQPKKIYNAKKDVVNTKMFEEELKSRSLDTLAEQVALLKEQALQTVCLAGTKVHHKC 60
Db	
QY	22 EPPTQPKKIYNAKKDVVNTKMFEEELKSRSLDTLAEQVALLKEQALQTVCLAGTKVHHKC 81
Db	
QY	61 FLAFTQTKTFHEASDCISRGGTLSTPQTGSENDALYEYLROSVGNBEAIWLGINDMAAE 120
Db	
QY	82 FLAFTQTKTFHEASDCISRGGTLSTPQTGSENDALYEYLROSVGNBEAIWLGINDMAAE 141
Db	
QY	121 GTWVDMTCARIAYKNWETETTAQPDGGKTENC AVLSGAANGKWFDKRCRDQLPYTCQFGI 180
Db	
QY	142 GTWVDMTCARIAYKNWETETTAQPDGGKTENC AVLSGAANGKWFDKRCRDQLPYTCQFGI 201
Db	
QY	181 V 181
Db	202 V 202

RESULT 4
US-09-602-877A-99

; Sequence 99, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-602-877A-99

Query Match 44.0%; Score 422.5; DB 4; Length 197;
Best Local Similarity 51.3%; Pred. No. 2.9e-40;
Matches 80; Conservative 27; Mismatches 46; Indels 3; Gaps 1;

QY 25 ELKSLDTLAORVALLKEQALQTVCLKGTQVHKMCFLAFTQTKTFHEASEDCISRGCTL 84
DB 43 DLKTOIEKLWTEVNALKIEQALQTVCLKGTQVHKMCFLAFTQTKTFHEASEDCISRGCTL 102
QY 85 STPTGSENDALEYLRQSVGNEABIWGLNDMAAEGTWDMTGARIAYKNWETEITAQP 144
DB 103 VIPRNSDEINADQYKSLPGVNDVFWLGINDMVTEGKFDVNGIAISFLNWD---RAQP 159
QY 145 DGGKTENCALVSGAANGKWFDRCDQLPYICQFGI 180
DB 160 NGGKRENCVLFQSAQKQKWSDEACRSKRYICEFTI 195

RESULT 5

US-09-065-040-12
; Sequence 12, Application US/09065040
; Patent No. 6541217
; GENERAL INFORMATION:
; APPLICANT: Hiraoka, Atsunobu
; APPLICANT: Sugimura, Atsushi
; APPLICANT: Mio, Hiroyuki
; TITLE OF INVENTION: HEMATOPOIETIC STEM CELL GROWTH FACTOR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; ADDRESSEE: DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,040
; FILING DATE: 27-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 262252/1996
; FILING DATE: 27-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 087242/1997
; FILING DATE: 24-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP97/02349
; FILING DATE: 07-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.

; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 04853.0026-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-040-12

Query Match 23.2%; Score 222.5; DB 4; Length 328;
Best Local Similarity 28.6%; Pred. No. 6.1e-17;
Matches 55; Conservative 33; Mismatches 73; Indels 31; Gaps 5;

QY 17 VVNTKMPF-----ELKSLDTLAORVALLKEQALQTV-----CLKGTQVHKMCFLA 63
DB 136 VLDTRVVELTQGLRRLRDAASDTRDSVQALKEVQVRSEQHGRLEGCLGLRLGHKCFLL 195
QY 64 FTQTKTFHEASEDCISRGCTLSTPQTGSENDALEYLRQSVG-NEABIWGLNDMAAEGT 122
DB 196 SRDFETQAAQAQRCVARGGSLAQPADROQMDALSRYLRAALAPYNWFWLGVHDDRSEGL 255
QY 123 WDMTGARIAYKNWETEITA-----QPDGKKTENCALVSGAANGKWFDR 166
DB 256 YLFENGQVRVFPFHWRLALSPESGAQPSAASHPLSPDQPNGGILENC-VAQASDDGSGWWDH 314
QY 167 RCDQLPYICQF 178
DB 315 DCERRLYFVCEP 326

RESULT 6

US-09-065-040-8
; Sequence 8, Application US/09065040
; Patent No. 6541217
; GENERAL INFORMATION:
; APPLICANT: Hiraoka, Atsunobu
; APPLICANT: Sugimura, Atsushi
; APPLICANT: Mio, Hiroyuki
; TITLE OF INVENTION: HEMATOPOIETIC STEM CELL GROWTH FACTOR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; ADDRESSEE: DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,040
; FILING DATE: 27-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 262252/1996
; FILING DATE: 27-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 087242/1997
; FILING DATE: 24-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP97/02349
; FILING DATE: 07-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.


```
QY 166 KRCRDLPYICOF 178
Db 309 HDCORRLYYVCEP 321

RESULT 9
US-09-513-999C-4342
; Sequence 4342, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4342
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: -21...-1
; OTHER INFORMATION: score 11.2
; OTHER INFORMATION: seq LILCLFSLLTQVTT/EP
US-09-513-999C-4342

Query Match 19.7%; Score 189; DB 4; Length 59;
Best Local Similarity 100.0%; Pred. No. 3.1e-14; Mismatches 0; Indels 0; Gaps 0;
Matches 38; Conservative 0;

QY 1 EPPTQPKKIVNAKDVNTKMFELKSRDLTLAQEVA 38
Db 22 EPPTQPKKIVNAKDVNTKMFELKSRDLTLAQEVA 59

RESULT 10
US-08-340-428B-49
; Sequence 49, Application US/08340428B
; Patent No. 5648465
; GENERAL INFORMATION:
; APPLICANT: MARGOLIS, Richard U.
; APPLICANT: RAUCH, Uwe
; APPLICANT: MARGOLIS, Renee K.
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
; TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Broadway and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,428B
; FILING DATE: 14 No. 5648465ember 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/922,911
; FILING DATE: 03 August 1992

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: Margolis-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197.
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-340-428B-49

Query Match 19.4%; Score 186.5; DB 1; Length 1257;
Best Local Similarity 32.7%; Pred. No. 6.5e-12; Mismatches 20; Indels 27; Gaps 8;
Matches 48; Conservative 20;

QY 49 VCLKGT-----KVMKCFLAFTQTKTFHEASEDCISRGGTLSTPQTGSENDALYEYL 100
Db 1021 LCEKDETCGDRGWHKFGQHCYRYFAHRAWEDEDCRRAGHLTSVHSPEHKFI---- 1076
QY 101 RQSVGNRAEIMGLNDMAARG--TWVDMTGARIAYKNWETEITAQPD-----GGKTENCAY 154
Db 1077 -NSFGHE-NSWIGLNDRTVERDFQWTDNTG--LQYENWREK---QPDNFFAGG--EDCVV 1127
QY 155 LSGAANGKWFDRKCRDQLPYICQGIIV 181
Db 1128 MVAHENGWRNDVPCNYNLPYVCKKGTIV 1154

RESULT 11
US-09-198-603C-2
; Sequence 2, Application US/09198603C
; Patent No. 6337193
; GENERAL INFORMATION:
; APPLICANT: TULLY, Raymond E.
; APPLICANT: CALTAGIRONE, G. Thomas
; APPLICANT: MOYER, Shawn S.
; APPLICANT: RONNING, Michael T.
; TITLE OF INVENTION: EXPRESSION OF MANNOSE-BINDING PROTEIN IN METHYLOTROPHIC
; TITLE OF INVENTION: YEAST
; FILE REFERENCE: A7290
; CURRENT APPLICATION NUMBER: US/09/198,603C
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Human
US-09-198-603C-2

Query Match 19.3%; Score 185.5; DB 3; Length 248;
Best Local Similarity 30.2%; Pred. No. 7.1e-13; Mismatches 28; Indels 21; Gaps 5;
Matches 45; Conservative 28;

QY 42 EQQALQTVCLK-----GTVKVMKCFLAFTQTKTFHEASEDCISRGGTLSTPQTGSE 92
Db 111 ERKALQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKKVLCVKFOASVATPRNAE 170
QY 93 NDALVEYLRQSVGNRAEIMGLNDMAARGTWVDMTGARIAYKNWETEITAQPD-DGSKTEN 151
Db 171 NGAIONLIKE-----EAPLGITDTEGQFVDLGNRLTYTNWN---EGEPNAGSDED 221
QY 152 CAVLSGAANGKWFDRKCRDQLPYICQFGI 180
Db 222 CVLL--LKNQWQNDVPCSTSHLAVCEPFI 248
```

RESULT 12

US-09-600-932-27
; Sequence 27, Application US/09600932
; Patent No. 6787639
; GENERAL INFORMATION:
; APPLICANT: Wakamiya, No. 6787639utaka
; TITLE OF INVENTION: NOVEL COLLECTIN
; FILE REFERENCE: 19036/36615
; CURRENT APPLICATION NUMBER: US/09/600,932
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: PCT/JP98/03328
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: mannan-binding protein (MBP)
US-09-600-932-27

Query Match 19.3%; Score 185.5; DB 4; Length 248;
Best Local Similarity 30.2%; Pred. No. 7.1e-13;
Matches 45; Conservative 28; Mismatches 55; Indels 21; Gaps 5;
QY 42 EQALQTVCLK-----GTVHMKCFIAFTQTKTFHEASEDCISRGGTLSTPTGSE 92
DB 111 ERKALQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKQASVATPRNAE 170
QY 93 NDALYELRQSVGNEAEIWLGLNDMAAEGTWDMTGARIAYKNWETEITAQP-DGGKTEN 151
DB 171 NGAIONLIKE-----EAFGLITDEKTEGQFVDLTGNRLTYTNW---EGEPNAGSDED 221
QY 152 CAVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 222 CVLL--LKNQWMDVPCSTSHLAVCEPFI 248

RESULT 13

US-09-949-016-5907
; Sequence 5907, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5907
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-5907

Query Match 19.3%; Score 185.5; DB 4; Length 248;
Best Local Similarity 30.2%; Pred. No. 7.1e-13;
Matches 45; Conservative 28; Mismatches 55; Indels 21; Gaps 5;
QY 42 EQALQTVCLK-----GTVHMKCFIAFTQTKTFHEASEDCISRGGTLSTPTGSE 92
DB 111 ERKALQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKQASVATPRNAE 170

QY 93 NDALYELRQSVGNEAEIWLGLNDMAAEGTWDMTGARIAYKNWETEITAQP-DGGKTEN 151
DB 171 NGAIONLIKE-----EAFGLITDEKTEGQFVDLTGNRLTYTNW---EGEPNAGSDED 221
QY 152 CAVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 222 CVLL--LKNQWMDVPCSTSHLAVCEPFI 248

RESULT 14

US-09-949-016-9370
; Sequence 9370, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9370
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9370

Query Match 19.3%; Score 185.5; DB 4; Length 263;
Best Local Similarity 30.2%; Pred. No. 7.8e-13;
Matches 45; Conservative 28; Mismatches 55; Indels 21; Gaps 5;
QY 42 EQALQTVCLK-----GTVHMKCFIAFTQTKTFHEASEDCISRGGTLSTPTGSE 92
DB 126 ERKALQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKQASVATPRNAE 185
QY 93 NDALYELRQSVGNEAEIWLGLNDMAAEGTWDMTGARIAYKNWETEITAQP-DGGKTEN 151
DB 186 NGAIONLIKE-----EAFGLITDEKTEGQFVDLTGNRLTYTNW---EGEPNAGSDED 236
QY 152 CAVLSGAANGKWFDRKCRDQLPYICQFGI 180
DB 237 CVLL--LKNQWMDVPCSTSHLAVCEPFI 263

RESULT 15

US-09-011-735-6
; Sequence 6, Application US/09011735B
; Patent No. 6110708
; GENERAL INFORMATION:
; APPLICANT: Wakamiya, No. 6110708utaka
; TITLE OF INVENTION: Recombinant Conglutinin and Producing Method Thereof
; FILE REFERENCE: 19036/34548
; CURRENT APPLICATION NUMBER: US/09/011,735B
; CURRENT FILING DATE: 1998-05-22
; EARLIER APPLICATION NUMBER: JP 7-209698
; EARLIER FILING DATE: 1995-08-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (2)

; OTHER INFORMATION: Xaa is a protein-constituting amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (3)
; OTHER INFORMATION: Xaa is a protein-constituting amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa is a protein-constituting amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (6)
; OTHER INFORMATION: Xaa is a protein-constituting amino acid
US-09-011-735-6

Query Match 18.6%; Score 178.5; DB 3; Length 161;
Best Local Similarity 31.7%; Pred. No. 2.3e-12;
Matches 40; Conservative 23; Mismatches 58; Indels 5; Gaps 2;
QY 53 GTKVHMKCFLEPTQTKTFHEASEDCISRGCTLTPTQGSNDALYEYLRQSVGNABEHWL 112
DB 41 GQAVGEKIFKTAGAVKYSYSDAEQLCREAKGQLASPRSSAENEAVTQMVR---AQEKNAYL 97
QY 113 GLNDMAAEGTAVDVTGARIAYKNWETEITAOQPDGKTCNCAVLSCAANGKWFDEKCRDQL 172
DB 98 SMDISTGRFTYPTGEILVYSNWADGEPNNSDEQPCNCEI--FPDGKNDVPCSKQL 155
QY 173 PYICQF 178
DB 156 LVICEF 161

Search completed: April 26, 2005, 11:54:46
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:53:11 ; Search time 176 Seconds
(without alignments)
526.627 Million cell updates/sec

Title: US-09-445-576A-7

Perfect score: 960

Sequence: 1 EPPTQKPKKIVNAKKDVNT.....KWFDRKCRDQLPYCQFGIV 181

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_eprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	100.0	202	1 TETN HUMAN	P05452 homo sapien
2	792	82.5	202	1 TETN MOUSE	P43025 mus musculus
3	792	82.5	202	2 Q8CFZ6	Q8cfz6 mus musculus
4	725	75.5	133	2 Q68DS3	Q68ds3 homo sapien
5	720.5	75.1	201	2 Q9DDA4	Q9dda4 gallus gall
6	455	47.4	193	2 Q68KU1	Q68kul xenopus lae
7	431.5	44.9	197	2 Q28008	Q28008 bos taurus
8	430.5	44.8	196	2 Q9EPW4	Q9epw4 mus musculus
9	422.5	44.0	197	1 CLB1 HUMAN	O75596 homo sapien
10	422.5	44.0	206	2 Q6UXF5	Q6uxf5 homo sapien
11	408	42.5	166	1 TETN CARSP	P26258 carcharhinu
12	222.5	23.2	328	1 SCGF_RAT	O88201 rattus norv
13	221.5	23.1	328	1 SCGF_MOUSE	O88200 mus musculus
14	220.5	23.0	323	1 SCGF_HUMAN	Q9Y240 homo sapien
15	207.5	21.6	369	1 PSPD_BOVIN	P35246 bos taurus
16	199.5	20.8	369	2 Q863A1	Q863a1 bos taurus
17	188.5	19.6	652	2 Q8NSX1	O8nsx1 dugesia tig
18	187.5	19.5	374	1 PSPD_MOUSE	P50404 mus musculus
19	186.5	19.4	244	1 MABC_MOUSE	P41317 mus musculus
20	186.5	19.4	1257	1 PGCN_RAT	P55067 rattus norv
21	185.5	19.3	248	1 MABC_HUMAN	P11226 homo sapien
22	185.5	19.3	248	1 MABC_HYLCO	Q66s58 hylobates c
23	184.5	19.2	248	1 MABC_GORGO	Q66s60 gorilla gor
24	183.5	19.1	248	1 MABC_PANTR	Q66s63 pan troglod
25	183.5	19.1	1290	2 Q9W6E1	Q9w6e1 gallus gall
26	183	19.1	246	2 Q919Q7	Q919q7 carassius a
27	182.5	19.0	248	1 MABC_PONPY	Q66s64 pongo pygma
28	182.5	19.0	371	1 CL46_BOVIN	Q8mh29 bos taurus
29	182.5	19.0	1268	1 PGCN_MOUSE	P55066 mus musculus
30	182.5	19.0	1268	2 Q6PIE3	Q6pie3 mus musculus
31	181.5	18.9	248	1 MABC_HYLLA	Q66s54 hylobates 1

RESULT 1

ID	TETN HUMAN	STANDARD;	PRT;	202 AA.
AC	P05452;			
DT	01-NOV-1998 (Rel. 09, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DE	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Tetranectin precursor (TN) (Plasminogen-kringle 4 binding protein).			
GN	Name=TNA;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=92365345; PubMed=1354271;			
RA	Wewer U.M., Albrechtsen R.;			
RT	"Tetranectin, a plasminogen kringle 4-binding protein. Cloning and			
RT	gene expression pattern in human colon cancer.";			
RL	Lab. Invest. 67:253-262(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92380263; PubMed=1511740; DOI=10.1016/0014-5793(92)80729-Z;			
RA	Berglund L., Petersen T.E.;			
RT	"The gene structure of tetranectin, a plasminogen binding protein.";			
RL	FEBS Lett. 309:15-19(1992).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[4]			
RP	SEQUENCE OF 22-202, AND VARIANTS SER-55 AND MET-58.			
RX	MEDLINE=88107595; PubMed=3427041;			

ALIGNMENTS

ID	TETN HUMAN	STANDARD;	PRT;	202 AA.
AC	P05452;			
DT	01-NOV-1998 (Rel. 09, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DE	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Tetranectin precursor (TN) (Plasminogen-kringle 4 binding protein).			
GN	Name=TNA;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=92365345; PubMed=1354271;			
RA	Wewer U.M., Albrechtsen R.;			
RT	"Tetranectin, a plasminogen kringle 4-binding protein. Cloning and			
RT	gene expression pattern in human colon cancer.";			
RL	Lab. Invest. 67:253-262(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92380263; PubMed=1511740; DOI=10.1016/0014-5793(92)80729-Z;			
RA	Berglund L., Petersen T.E.;			
RT	"The gene structure of tetranectin, a plasminogen binding protein.";			
RL	FEBS Lett. 309:15-19(1992).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[4]			
RP	SEQUENCE OF 22-202, AND VARIANTS SER-55 AND MET-58.			
RX	MEDLINE=88107595; PubMed=3427041;			

Q61361	mus musculus
Q80wt7	mus musculus
Q8axr8	anguilla ja
Q66s50	macaca fasc
Q66s65	papio papio
Q28518	macaca mula
P23805	bos taurus
P08661	rattus norv
O14594	homo sapien
Q63097	rattus norv
Q84t3	rattus norv
Q6az45	rattus norv
Q28062	bos taurus
Q28517	macaca mula

RA Fuhlendorff J., Clemmensen I., Magnusson S.;
RT "Primary structure of tetranectin, a plasminogen kringle 4 binding
RT plasma protein: homology with asialoglycoprotein receptors and
RT cartilage proteoglycan core protein.";
RL Biochemistry 26:6757-6764 (1987).
RN [5]
RN SEQUENCE OF 1-36 FROM N.A.
RP TISSUE=Placenta;
RC Sorensen C.B., Berglund L., Petersen T.E.;
RT "Cloning and mapping of the murine tetranectin gene.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [6]
RN PARTIAL SEQUENCE, MASS SPECTROMETRY, AND VARIANT GLY-106.
RP TISSUE=Plasma;
RC MEDLINE=20080486; PubMed=10614823;
RA Jaquinod M., Holtet T.L., Ezerodt M., Clemmensen I., Thøgersen H.C.,
RA Roepstorff P.;
RT "Mass spectrometric characterisation of post-translational
RT modification and genetic variation in human tetranectin.";
RL Biol. Chem. 380:1307-1314 (1999).
RN [7]
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RP MEDLINE=97398360; PubMed=9256258; DOI=10.1016/S0014-5793(97)00664-9;
RA Nielsen B.B., Kasstrup J.S., Rasmussen H., Holtet T.L., Graversen J.H.,
RA Ezerodt M., Thøgersen H.C., Larsen I.K.;
RT "Crystal structure of tetranectin, a trimeric plasminogen-binding
RT protein with an alpha-helical coiled coil.";
RL FEBS Lett. 412:388-396 (1997).
RN [8]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 66-202.
RP MEDLINE=98437604; PubMed=9757090; DOI=10.1107/S0907444997016806;
RA Kasstrup J.S., Nielsen B.B., Rasmussen H., Holtet T.L., Graversen J.H.,
RA Ezerodt M., Thøgersen H.C., Larsen I.K.;
RT "Structure of the C-type lectin carbohydrate recognition domain of
RT human tetranectin.";
RL Acta Crystallogr. D 54:757-766 (1998).
CC -I- FUNCTION: Tetranectin binds to plasminogen and to isolated kringle
CC 4. May be involved in the packaging of molecules destined for
CC exocytosis.
CC -I- SUBUNIT: Homotrimer.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Found in plasma.
CC -I- MASS SPECTROMETRY: MW=20535.8; MW_ERR=2.4; METHOD=Electrospray;
CC RANGE=22-202; NOTE=Ref.6.
CC -I- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70910; CAA50265.1; -
DR EMBL; X70911; CAA50265.1; JOINED.
DR EMBL; X70912; CAA50265.1; JOINED.
DR EMBL; X64559; CAA45860.1; -
DR EMBL; BC011024; AAH11024.1; -
DR EMBL; X98121; CAA66803.1; -
DR FIR; S24126; ITHUN.
DR PDB; 1HTN; X-ray; @=-.
DR PDB; 1TN3; X-ray; @=66-202.
DR Genew; HGNC:11891; TNA.
DR H-InvDB; HIX0003237; -
DR MIM; 187520; -
DR GO; GO:0001501; P:skeletal development; TAS.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; Lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT. 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.

DR PROSITE; PS0041; C_TYPE_LECTIN_2; 1.
KW 3D-structure; Direct protein sequencing; Glycoprotein; Lectin; Plasma;
KW Polymorphism; Signal.
FT SIGNAL 1 21
FT CHAIN 22 202
FT DOMAIN 77 198
FT DISULFID 71 81
FT DISULFID 98 197
FT DISULFID 173 189
FT CARBOHYD 25 25
FT VARIANT 55 55
FT O-linked (Galnac. . .).
FT A -> S.
FT /FTID=VAR_004189.
FT V -> M.
FT /FTID=VAR_004190.
FT S -> G (in dbSNP:13963).
FT /FTID=VAR_012318.
FT
FT STRAND 70 73
FT STRAND 79 89
FT HELIX 91 100
FT TURN 101 102
FT STRAND 104 105
FT HELIX 111 124
FT TURN 125 125
FT TURN 127 128
FT STRAND 130 140
FT TURN 141 142
FT STRAND 143 146
FT TURN 147 148
FT STRAND 151 151
FT STRAND 157 157
FT TURN 160 162
FT HELIX 168 170
FT STRAND 173 177
FT TURN 178 182
FT STRAND 183 187
FT TURN 189 190
FT STRAND 193 200
SQ SEQUENCE 202 AA; 22567 MW; 2B0DCB5DF22E1AB8 CRC64;

Query Match 100.0%; Score 960; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.8e-77;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPPTQPKPKIVNAKDVVNTKMFELKSRDLTLAQEVALLEKQALQTVCLKGVHMKC 60
DB 22 EPPTQPKPKIVNAKDVVNTKMFELKSRDLTLAQEVALLEKQALQTVCLKGVHMKC 81
QY 61 FLAFTQTKTTFHEASEDCISRGGTLSTPTGSENDALEYLRQSVGNBAEIIWGLNDMAAE 120
DB 82 FLAFTQTKTTFHEASEDCISRGGTLSTPTGSENDALEYLRQSVGNBAEIIWGLNDMAAE 141
QY 121 GTWVDMTGARIYKNWETETITAOQPDGKKTENCAVLSGAANGKWFEDKRCRDOLPYICQFI 180
DB 142 GTWVDMTGARIYKNWETETITAOQPDGKKTENCAVLSGAANGKWFEDKRCRDOLPYICQFI 201

QY 181 V 181
DB 202 V 202

RESULT 2
TETN_MOUSE
ID TETN_MOUSE STANDARD; PRT; 202 AA.
AC P43025;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Tetranectin precursor (TN) (Plasminogen-kringle 4 binding protein).
GN Name=Tna;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RA MEDLINE=9513796; PubMed=7835708; DOI=10.1016/0378-1119(94)00703-U;
RX Soerensen C.B., Berglund L., Petersen T.E.;
RT "Cloning of a cDNA encoding murine tetranectin.";
RL Gene 152:243-245(1995).
RN RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=96116955; PubMed=8563165;
RA Ibaraki K., Kozak C.A., Wewer U.M., Albrechtsen R., Young M.F.;
RT "Mouse tetranectin: cDNA sequence, tissue-specific expression, and
chromosomal mapping.";
RL Mamm. Genome 6:693-696(1995).
RN RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=98072445; PubMed=9409787; DOI=10.1016/S0378-1119(97)00451-4;
RA Soerensen C.B., Berglund L., Petersen T.E.;
RT "Cloning of the murine tetranectin gene and 5'-flanking region.";
RL Gene 201:199-202(1997).
CC -!- FUNCTION: Tetranectin binds to plasminogen and to isolated kringle
exocytosis (By similarity).
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Highest expression in lung and skeletal
muscle.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; X79199; CAA55791.1; -
DR EMBL; U08595; AAA96811.1; -
DR EMBL; X98122; CAA66804.1; -
DR PIR; JC4031; JC4031.
DR HSP; P05452; ITN3.
DR MGD; MGI:104540; Tna.
DR GO; GO:0001501; P:skeletal development; IMP.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; Lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
KW Lectin; Plasma; Signal.
FT SIGNAL 1 21 By similarity.
FT CHAIN 22 202 Tetranectin.
FT DOMAIN 77 198 C-type lectin.
FT DISULFID 71 81 By similarity.
FT DISULFID 98 197 By similarity.
FT DISULFID 173 189 By similarity.
FT CONFLICT 19 20 LT -> VI (in Ref. 2).
FT CONFLICT 84 84 A -> T (in Ref. 2).
FT CONFLICT 180 180 A -> R (in Ref. 2).
FT CONFLICT 188 188 R -> Q (in Ref. 2).
SQ SEQUENCE 202 AA; 22257 MW; 639E7334D5E8E04E CRC64;
Query Match 82.5%; Score 792; DB 1; Length 202;
Best Local Similarity 80.7%; Pred. No. 1.6e-62;
Matches 146; Conservative 14; Mismatches 21; Indels 0; Gaps 0;
QY 1 EPPTOKPKIVNAKDVNTKMFELKSRDLTAQEVALLKEQOALQTVCLKGTVMKLC 60
DB 22 ESPTPKAKAANKAKDLVSSKMFELKSRDMVLAQEVALLKEQALQTVCLKGTVMKLC 81

```

```

QY 61 FLAFTQTKTFHEASDCISRGCTLTSTPQTSNDALYLYLRSQVGNENAEIWLGLNDMAAE 120
DB 82 LLAFATQPTTFHEASDCISQGGTLGTPTQSELENEALFEYARHSVGNNDANIWLGLNDMAAE 141
QY 121 GTWVDMTCARITAYKNWETETTAQPDGGKTKENCANVLSCGAANGKWFDKRCRDOLPYICQFGI 180
DB 142 GAWVDMTCGLLAYKNWETETITTPDGGKAENCAALSGAANGKWFDKRCRDOLPYICQFAI 201
QY 181 V 181
DB 202 V 202
RESULT 3
Q8CFZ6 PRELIMINARY; PRT; 202 AA.
AC Q8CFZ6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tetranectin (Plasminogen binding protein).
GN Name=Tna;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=2338257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughlano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035043; AAH35043.1; -
DR HSP; P05452; ITN3.
DR MGD; MGI:104540; Tna.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); TAS.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR GO; GO:0005529; F:sugar binding; TAS.
DR GO; GO:0001501; P:skeletal development; IMP.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; Lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 202 AA; 22255 MW; 65707A362F93A3C3 CRC64;

```

```

Query Match      82.5%; Score 792; DB 2; Length 202;
Best Local Similarity 80.7%; Pred. No. 1.6e-62;
Matches 146; Conservative 14; Mismatches 21; Indels 0; Gaps 0;

QY 1 EPTQPKKIVNAKDVVNTKMFELKSRDLTIAQEVALLKEQALQTVCLKGTVMKCF 60
D 22 ESPTPKAKAANKKOLVSSKMFELKSRDLTIAQEVALLKEQALQTVCLKGTVMKCF 81
QY 61 FLAFTQPKTTFHEASEDCISRGGLTSTPQTGSENDALYELRQSVGNEAEIWLGLNDMAAE 120
D 82 LLAFTQPKTTFHEASEDCISRGGLTSTPQTGSENDALYELRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWDMTGARIAYKNWETETIAQPDGKTCNCVLSGAANGKWFDRCDQLPYICQFGI 180
D 142 GAWWDMTGLLAYKNWETETITQPDGKTCNCVLSGAANGKWFDRCDQLPYICQFAI 201
QY 181 V 181
D 202 V 202

RESULT 4
Q68DS3 PRELIMINARY; PRT; 133 AA.
ID Q68DS3
AC Q68DS3
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein DKF2p686H17246 (Fragment).
GN Name=DKF2p686H17246;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Salivary gland;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Meves H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL: CR749292; CAH18147.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C-TYPE LECTIN 2; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 133 AA; 14708 MW; EFBA9A42EFE341F4 CRC64;

Query Match      75.5%; Score 725; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 9.4e-57;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VCLKGTVMKCFLAFTQPKTTFHEASEDCISRGGLTSTPQTGSENDALYELRQSVGNEA 108
D 1 VCLKGTVMKCFLAFTQPKTTFHEASEDCISRGGLTSTPQTGSENDALYELRQSVGNEA 60
QY 109 EIWLGLNDMAAGTWDMTGARIAYKNWETETIAQPDGKTCNCVLSGAANGKWFDRKRC 168
D 61 EIWLGLNDMAAGTWDMTGARIAYKNWETETIAQPDGKTCNCVLSGAANGKWFDRKRC 120
QY 169 RDQLPYICQFGIV 181
D 121 RDQLPYICQFGIV 133

RESULT 5
Q9DD4 PRELIMINARY; PRT; 201 AA.
ID Q9DD4
AC Q9DD4
DT 01-WAR-2001 (TrEMBLrel. 16, Created)
DT 01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tetranectin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Wewer U.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL: AJ277116; CAC20217.1; -.
DR HSP; P05452; ITN3.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pncreatis_ac.
DR Pfam; PF00059; Lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C-TYPE LECTIN 2; 1.
SQ SEQUENCE 201 AA; 22172 MW; 7C7F235D24426AE8 CRC64;

Query Match      75.1%; Score 720.5; DB 2; Length 201;
Best Local Similarity 73.4%; Pred. No. 3.8e-56;
Matches 130; Conservative 23; Mismatches 21; Indels 3; Gaps 1;

QY 5 QPKKIVNAKDVVNTKMFELKSRDLTIAQEVALLKEQALQTVCLKGTVMKCF 64
D 28 QKP---AASKKGVSLKMIEDLKAMIDNISQEVALLKEQALQTVCLKGTVMKCF 84
QY 65 TQTKTFHEASEDCISRGGLTSTPQTGSENDALYELRQSVGNEAEIWLGLNDMAAGTW 124
D 85 SESKTYHEASEHCISQGGTLTPOGGEENDALYDMRKSIQNEAEIWLGLNDMAAGTW 144
QY 125 DMTGARIAYKNWETETIAQPDGKTCNCVLSGAANGKWFDRCDQLPYICQFGIV 181
D 145 DMTGSPTRYKNWETETITQPDGKTCNCVLSGAANGKWFDRCDQLPYICQFGIV 201

RESULT 6
Q66KU1 PRELIMINARY; PRT; 193 AA.
ID Q66KU1
AC Q66KU1
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE MGC85438 protein.
GN Name=MGC85438;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Whole;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smalish D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC078559; AAH78559.1; -;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 SQ SEQUENCE 193 AA; 22097 MW; 2FD58EF734E626D7 CRC64;

Query Match 47.4%; Score 455; DB 2; Length 193;
 Best Local Similarity 49.1%; Pred. No. 1.7e-32;
 Matches 86; Conservative 34; Mismatches 49; Indels 6; Gaps 2;

QY 4 TQPKKIVNAKDVVNTKMFELKSLRLDTLAEVALLKEQQAQTVCCLKGTVMKCFPLA 63
 DB 22 TUSAK--LXTQDHRKSEKDGKLTQIDKLRWELNSLKEMQALQTVCLGKTKHKCYLS 79

QY 64 FTQTKTFHEASDCISRGTTSTPQTSNDALYELROSVDNEAIEIWLGLNDMAAEGTW 123
 DB 80 FEETKHFHEANEDCIAKGGTLAI PRDSENNALRDYKKSLSHGSEFWLGLNDMNEGKF 139

QY 124 VDMTGARTAYKWETEITAPQGGKTENCANVLSGAANGKWFDRCDQLPYICQF 178
 DB 140 VDVNGVAITYFWNE----RIPKGRKKNCALLNQASQGWVDVCRSLKXYICEF 190

RESULT 7
 Q28008 ID Q28008 PRELIMINARY; PRT; 197 AA.
 AC Q28008
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-WAR-2004 (TREMBlrel. 26, Last annotation update)
 DE C-type lectin homolog precursor.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cartilage;
 RA Neame P.J., Boynton R.E.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U22298; AAC18614.1; -;
 DR HSP; P05452; ITN3.
 DR GO; GO:0005529; F:sugar binding; IEA.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam; PF00059; Lectin_C; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

KW Lectin; Signal.
 FT SIGNAL 1 24 Potential.
 FT CHAIN 25 197 C-type lectin homolog.
 SQ SEQUENCE 197 AA; 22215 MW; AAAC4280F41ACOF4 CRC64;

Query Match 44.9%; Score 431.5; DB 2; Length 197;
 Best Local Similarity 51.9%; Pred. No. 2.1e-30;
 Matches 81; Conservative 29; Mismatches 43; Indels 3; Gaps 1;

QY 25 ELKSLRLDTLAEVALLKEQQAQTVCCLKGTVMKCFPLAFTQTKTFHEASDCISRGTT 84
 DB 43 DLKTOVEKLREVNALKEMQALQTVCLGTRFKHKCYLAEGLEKHFHEANEDCISKGTT 102

QY 85 STPQTSNDALYELROSVDNEAIEIWLGLNDMAAEGTWDMTGARTAYKWETEITAPQ 144
 DB 103 VVPSRADSINALRDYKKSLSFGVNDFWLGLNDMVAEGKFDVINGLAISFLNWD---QAQ 159

QY 145 DGGKTENCANVLSGAANGKWFDRCDQLPYICQF 180
 DB 160 NGGKRENCALFSQSAQKWSDEACHSKRYICEFT 195

RESULT 8
 Q3EPW4 ID Q3EPW4 PRELIMINARY; PRT; 196 AA.
 AC Q3EPW4
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE C-type lectin superfamily 1.
 GN Name=Clectsl;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ES129/SV;
 RA Neame P.J., Grimm D.R.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317204; AAG48620.1; -;
 DR HSP; P05452; IHTN.
 DR MGD; MGI:2685642; Clectsl.
 DR GO; GO:0005529; F:sugar binding; IEA.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam; PF00059; Lectin_C; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 KW Lectin.
 SQ SEQUENCE 196 AA; 22191 MW; 9AB4C809D119E852 CRC64;

Query Match 44.8%; Score 430.5; DB 2; Length 196;
 Best Local Similarity 48.5%; Pred. No. 2.6e-30;
 Matches 83; Conservative 30; Mismatches 47; Indels 11; Gaps 2;

QY 8 KKVNAKDVVNTKMFELKSLRLDTLAEVALLKEQQAQTVCCLKGTVMKCFPLAFTQ 67
 DB 34 KRVKAKDD-----DLKSQVEKLREVNALKEMQALQTVCLGTRFKHKCYLAEG 85

QY 68 KTFHEASDCISRGTTSTPQTSNDALYELROSVDNEAIEIWLGLNDMAAEGTWDMT 127
 DB 86 KHYHEANEDCISKGTTLVVPRNSDEINALRDYKKSLSFGVNDFWLGLNDMVAEGKFDV 145

QY 128 GARTAYKWETEITAPQGGKTENCANVLSGAANGKWFDRCDQLPYICQF 178
 DB 146 GFVAVSFLNWD---RAQPSGGRKRENCALFSQSAQKWSDEACHSKRYICEF 193

RESULT 9
 CLEI_HUMAN ID CLEI_HUMAN STANDARD; PRT; 197 AA.
 AC 075596;

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE C-type lectin superfamily member 1 precursor (Cartilage-derived C-type
 DE lectin).
 GN Name=CLECSF1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cartilage;
 RX MEDLINE=99453719; PubMed=105241194; DOI=10.1016/S0167-4781(99)00087-1;
 RA Neame P.J., Tapp H., Grimm D.R.;
 RA "The cartilage-derived, C-type lectin (CLECSF1): structure of the gene
 RT and chromosomal location.";
 RL Biochim. Biophys. Acta 1446:193-202 (1999).
 CC -1- TISSUE SPECIFICITY: Restricted to cartilage.
 CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF077345; AAD12542.1; --
 DR EMBL; AF077344; AAD12542.1; JOINED.
 DR HSSP; P05452; IHTN.
 DR GENE; HGNC:2052; CLECSF1.
 DR GO; GO:0005529; F:sugar binding; TAS.
 DR GO; GO:0001501; P:skeletal development; TAS.
 DR InterPro; IPR001304; Lectin_C.
 DR Pfam; PF00059; Lectin_C; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C TYPE LECTIN 1; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 KW Lectin; Signal.
 FT SIGNAL 1 22 Potential.
 FT CHAIN 23 197 C-type lectin superfamily member 1.
 FT DOMAIN 74 192 C-type lectin.
 FT DISULFID 68 78 By similarity.
 FT DISULFID 95 191 By similarity.
 FT DISULFID 167 183 By similarity.
 SQ SEQUENCE 197 AA; BB924DDB7729A4 CRC64;
 Query Match 44.0%; Score 422.5; DB 1; Length 197;
 Best Local Similarity 51.3%; Pred. No. 1.3e-29;
 Matches 80; Conservative 27; Mismatches 46; Indels 3; Gaps 1;
 QY 25 ELKSLDITLAQEVALLKEQALQTVCLKGTVMKCFIAFTQTKTFHEASEDCISRGITL 84
 Db 43 DLKQIEKLWTEVNALEIQALQTVCLRGTVHKVCYLASEGLKHFHEANEDCISKGIL 102
 QY 85 STPQTGSNDALYELRQSGVNEAEIWLGLNDMAEGTWDMTGARIYKNWETETIAQP 144
 Db 103 VIPRNSDEINALQDYGRSLPGVNDPWLGLNDMTVEGKFDVNGIAISFLNWD---RAQP 159
 QY 145 DGGKTENCVAISGAANGKWFDRDQLPYICQFGI 180
 Db 160 NGGKRENCVLFSQAQGSDEACRSKRYICEFTI 195
 RESULT 10
 Q6UXF5 PRELIMINARY; PRT; 206 AA.
 ID Q6UXF5
 AC Q6UXF5
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE CLECSF1.
 GN ORFNames=UNQ700;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heldens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.H., Yansura D.,
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
 RA Godowski P.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270 (2003).
 DR EMBL; AY358376; AAQ88742.1; --
 DR HSSP; P35247; 1B08.
 DR GO; GO:0005529; F:sugar binding; IEA.
 DR InterPro; IPR001304; Lectin_C.
 DR Pfam; PF00059; Lectin_C; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C TYPE LECTIN 1; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 SQ SEQUENCE 206 AA; 23190 MW; F40F52210982F4C0 CRC64;
 Query Match 44.0%; Score 422.5; DB 2; Length 206;
 Best Local Similarity 51.3%; Pred. No. 1.4e-29;
 Matches 80; Conservative 27; Mismatches 46; Indels 3; Gaps 1;
 QY 25 ELKSLDITLAQEVALLKEQALQTVCLKGTVMKCFIAFTQTKTFHEASEDCISRGITL 84
 Db 52 DLKQIEKLWTEVNALEIQALQTVCLRGTVHKVCYLASEGLKHFHEANEDCISKGIL 111
 QY 85 STPQTGSNDALYELRQSGVNEAEIWLGLNDMAEGTWDMTGARIYKNWETETIAQP 144
 Db 112 VIPRNSDEINALQDYGRSLPGVNDPWLGLNDMTVEGKFDVNGIAISFLNWD---RAQP 168
 QY 145 DGGKTENCVAISGAANGKWFDRDQLPYICQFGI 180
 Db 169 NGGKRENCVLFSQAQGSDEACRSKRYICEFTI 204
 RESULT 11
 TETN CARSP STANDARD; PRT; 166 AA.
 ID TETN CARSP
 AC P26258;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Tetranectin-like protein.
 OS Carcharhinus springeri (Reef shark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
 OC Carcharhinidae; Carcharhinus.
 OX NCBI_TaxID=7809;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Neural arch cartilage;
 RX MEDLINE=93284081; PubMed=1304877;
 RA Neame P.J., Young C.N., Treep J.T.;
 RT "Primary structure of a protein isolated from reef shark (Carcharhinus
 RT springeri) cartilage that is similar to the mammalian C-type lectin
 RT homolog, tetranectin.";
 RL Protein Sci. 1:161-168 (1992).
 CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
 CC PIR; A37289; A37289.

```

DR HSSP; P05452; 1HTN.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; Lectin C; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C-TYPE LECTIN 2; 1.
KW Direct protein sequencing; Lectin.
FT DOMAIN 43 161 C-type lectin.
FT DISULFID 37 47
FT DISULFID 64 160
FT DISULFID 136 152
SQ SEQUENCE 166 AA; 18432 MW; 53BF812DEA5C6119 CRC64;

Query Match 42.5%; Score 408; DB 1; Length 166;
Best Local Similarity 44.5%; Pred. No. 2.1e-28;
Matches 77; Conservative 31; Mismatches 53; Indels 12; Gaps 2;

QY 6 KPKIVNAKQDVVNTKMPFELKSRDLTLAQEVALLEKQALQTVCLKGTQVHKMCKFLAFT 65
DB 2 KPSKSGKGD-----LNEDLKLWREVNSLKEMQALQTVCLKGTQVHKMCKFLAFT 52
QY 66 QTKTFHEASEDCISRGGTLSTPTQSGENDALYEYLRSQVGNBAEILWGLNDMAAEGTWYD 125
DB 53 GSKSYHAANEDCIAQGGTSLIPRSSDEGNSLSYAKSLVGARDFWIGVNDWTTGKFD 112
QY 126 MTGARIYKNWETETAPQDGGKTECNCAVLGAANGKWFKRCRDLPIYICOF 178
DB 113 VNGLPITTFNWD---RSKPVGGTRENCAVASTSGQGWSDDDVCRSEKRYICEY 162

RESULT 12
SCGF RAT STANDARD; PRT; 328 AA.
AC O88201;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Stem cell growth factor precursor (Lymphocyte secreted C-type lectin).
GN Name=Scgf;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone;
RX MEDLINE=98381038; PubMed=9705843; DOI=10.1006/bbrc.1998.9073;
RA Mio H., Kagami N., Yokokawa S., Kawai H., Nakagawa S., Takeuchi K.,
RA Sekine S., Hiraoka A.;
RT "Isolation and characterization of a cDNA for human, mouse, and rat
RT full-length stem cell growth factor, a new member of C-type lectin
RT superfamily."
RL Biochem. Biophys. Res. Commun. 249:124-130(1998).
CC !- FUNCTION: Stimulates the proliferation and differentiation of
CC hematopoietic precursor cells from various lineages, including
CC erythrocytes, lymphocytes, granulocytes and macrophages. Acts
CC synergistically with other cytokines, including IL-3, G-CSF, GM-CSF
CC and Flt3 ligand. Suppresses SCF-stimulated erythrocyte
CC proliferation (By similarity).
CC !- SUBCELLULAR LOCATION: Cytoplasmic and secreted (By similarity).
CC !- PTM: O-glycosylated. Probably sulfated on the O-glycans (By
CC similarity).
CC !- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB009246; BAA32406.1; -.
DR HSSP; P05452; 1HTN.

```

```

DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0008083; F:growth factor activity; IDA.
DR GO; GO:0005529; F:sugar binding; NAS.
DR GO; GO:0008284; P:positive regulation of cell proliferation; IDA.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00059; Lectin C; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C-TYPE LECTIN 2; 1.
KW Glycoprotein; Growth factor; Lectin; Signal.
FT SIGNAL 1 21 By similarity.
FT CHAIN 22 328 Stem cell growth factor.
FT DOMAIN 188 325 C-type lectin.
SQ SEQUENCE 328 AA; 36387 MW; C115188AC23D6257 CRC64;

Query Match 23.2%; Score 222.5; DB 1; Length 328;
Best Local Similarity 28.6%; Pred. No. 1.6e-11;
Matches 55; Conservative 33; Mismatches 73; Indels 31; Gaps 5;

QY 17 VVNTKMF-----ELKSLRLTLAQEVALLEKQALQTV-----CLKGTQVHKMCKFLA 63
DB 136 VLDTRVVELTQGLRLRLDAASDTRDSVQALKEVQVRSQEHGRLEGCLKGLRLGHKCFLL 195
QY 64 FTQYTFHEASEDCISRGGTLSTPTQSGENDALYEYLRSQVGNBAEILWGLNDMAAEGT 122
DB 196 SRDFETQAAQAARCKARGGSLAQPADRQOMDALSRYLAALAPYNWVWLVGHVDRRSEGL 255
QY 123 VWDMTGARIYKNWETETAPQDGGKTECNCAVLGAANGKWFKRCRDLPIYICOF 166
DB 256 YLFENGQVRFVFAHRAUSPESGAQPSAASHPLSPDQNGGILENC-VAQASDDGSWDH 314
QY 167 RCRDLPIYICOF 178
DB 315 DCERLYFVCFR 326

RESULT 13
SCGF MOUSE STANDARD; PRT; 328 AA.
AC O88200; O8C946; Q9CTF0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Stem cell growth factor precursor (Lymphocyte secreted C-type lectin).
GN Name=Scgf;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Calvaria;
RX MEDLINE=98381038; PubMed=9705843; DOI=10.1006/bbrc.1998.9073;
RA Mio H., Kagami N., Yokokawa S., Kawai H., Nakagawa S., Takeuchi K.,
RA Sekine S., Hiraoka A.;
RT "Isolation and characterization of a cDNA for human, mouse, and rat
RT full-length stem cell growth factor, a new member of C-type lectin
RT superfamily."
RL Biochem. Biophys. Res. Commun. 249:124-130(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Atschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

DR	InterPro; IPR001304; Lectin C.
DR	InterPro; IPR008994; Nucleic_acid_OB.
DR	Pfam; PF00059; Lectin C; 1.
DR	SMART; SM00034; CLECT; 1.
DR	PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR	PROSITE; PS00641; C-TYPE_LECTIN_2; 1.
KW	Glycoprotein; Growth factor; Lectin; Signal.
FT	SIGNAL 1 21 By similarity.
FT	CHAIN 22 328 Stem cell growth factor.
FT	DOMAIN 188 325 C-type lectin.
FT	CONFLICT 114 117 TYIL -> PFTV (in Ref. 3).
SQ	SEQUENCE 328 AA; 36451 MW; 309CI7A861EE135C CRC64;
Query Match	
Best Local Similarity 23.1%; Score 221.5; DB 1; Length 328;	
Matches 55; Conservative 36; Mismatches 31; Gaps	
Qy	11 VNAKDVVNTKPE-----ELKRLDTLAEVALLKEQQ-----ALQTVCLKGTKVH 57
Db	130 LHLVHLVLDTRVVELTQGLRQDAASDTRDSVQALKEQEHGRLEGCLGLRLG 189
Qy	58 MKCFLLATQTQTFHEASEDCISRGSTLTPTGSENDALYEVLRSVG-NEAEIWLGLND 116
Db	190 HKCEFLSRDPFTQAAQARCAKGSLSQAQPADRQMDALSRYLRALAPYNNPVMVLGVD 249
Qy	117 MAASEGTWVDMTGARIAYKNWETEITA-----QPDGKGTENCACVLSGAAN 160
Db	250 RRSEGLYLFENGQVSPFAWHRAFSLESGAQSAAATHPLSPDQFNGGVLENC-VAQASDD 308
Qy	161 GKWFDKRCRDQLPYICQF 178
Db	309 GSNWHDHCERRLYFVCFE 326
RESULT 14	
SCGF	HUMAN
ID	-SCGF HUMAN STANDARD; PRT; 323 AA.
AC	Q9Y240;
DT	10-OCT-2003 (Rel. 42, Created)
DT	10-OCT-2003 (Rel. 42, Last sequence update)
DT	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Stem cell growth factor precursor (Lymphocyte secreted C-type lectin)
DE	(p47).
GN	Name=SCGF; Synonyms=LSLCL;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A., SEQUENCE OF 22-39; 229-243 AND 314-323, TISSUE
RP	SPECIFICITY, AND O-GLYCOSYLATION.
RC	TISSUE=Bone marrow;
RX	MEDLINE=98113146; PubMed=9442024; DOI=10.1074/jbc.273.4.1911;
RA	Bannwarth S., Gordanengo V., Lesimple J., Lefebvre J.-C.;
RT	"Molecular cloning of a new secreted sulfated mucin-like protein with
RT	a C-type lectin domain that is expressed in lymphoblastic cells.";
RJ	J. Biol. Chem. 273:1911-1916(1998).
RN	[2]
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC	TISSUE=Bone marrow;
RX	MEDLINE=98381038; PubMed=9705843; DOI=10.1006/bbrc.1998.9073;
RA	Mio H., Kagami N., Yokokawa S., Kawai H., Nakagawa S., Takeuchi K.,
RA	Seikine S., Hiraoka A.;
RT	"Isolation and characterization of a cDNA for human, mouse, and rat
RT	full-length stem cell growth factor, a new member of C-type lectin
RT	superfamily.";
RL	Biochem. Biophys. Res. Commun. 249:124-130(1998).
RN	[3]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99216429; PubMed=10198175; DOI=10.1006/geno.1999.5762;
RA	Bannwarth S., Gordanengo V., Grosgeorge J., Turc-Carel C.,
RA	Lefebvre J.-C.;
RT	"Cloning, mapping, and genomic organization of the LSLCL gene,

RT encoding a new lymphocytic secreted mucin-like protein with a C-type
 RT lectin domain: a new model of exon shuffling.";
 RL Genomics 57:316-317(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Skin;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahay J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP FUNCTION, AND TISSUE SPECIFICITY.
 RX MEDLINE=2197201; PubMed=11920266; DOI=10.1038/sj/thj/6200118;
 RA Hirooka A., Yano K.-I., Kagami N., Takeshige K., Mio H., Anazawa H.,
 RA Sugimoto S.;
 RT "Stem cell growth factor: in situ hybridization analysis on the gene
 RT expression, molecular characterization and in vitro proliferative
 RT activity of a recombinant preparation on primitive hematopoietic
 RT progenitor cells.";
 RL Hematology J. 2:307-315(2001).
 RN [6]
 RP SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RX MEDLINE=21663069; PubMed=11803813; DOI=10.1016/S0764-4469(01)01392-0;
 RA Perrin C., Bayle J., Bannwarth S., Michiels J.-F., Heudier P.,
 RA Lefebvre J.-C., Giordanengo V.;
 RT "Expression of LSCL, a new C-type lectin, is closely restricted, in
 RT bone marrow, to immature neutrophils.";
 RL C. R. Acad. Sci. III, Sci. Vie 324:1125-1132(2001).
 CC -!- FUNCTION: Stimulates the proliferation and differentiation of
 CC hematopoietic precursor cells from various lineages, including
 CC erythrocytes, lymphocytes, granulocytes and macrophages. Acts
 CC synergistically with other cytokines, including IL-3, G-CSF, GM-CSF
 CC and FLT3 ligand. Suppresses SCF-stimulated erythrocyte
 CC proliferation.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and secreted.
 CC -!- TISSUE SPECIFICITY: Expressed in skeletal tissues including bone
 CC marrow, chondrocytes, primary ossification center-associated
 CC cells, the perichondrium and periosteum. Lower levels of
 CC expression were detected in spleen, thymus, appendix and fetal
 CC liver.
 CC -!- DEVELOPMENTAL STAGE: In the bone marrow, expression is limited to
 CC immature neutrophils. Expression was not detected in circulating
 CC mature neutrophils.
 CC -!- PTM: O-glycosylated. Probably sulfated on the O-glycans.
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF020044; AAC39569.1; -;
 DR EMBL; AB009244; BAA32404.1; -;

DR EMBL; AF087659; AAD26533.1; -;
 DR EMBL; BC005810; AAH05810.1; -;
 DR HSSP; P05452; 1HTN.
 DR Genew; HGNC:10576; SCGF.
 DR MIM; 604713; -;
 DR GO; GO:0005576; C:extracellular; IDA.
 DR GO; GO:0008093; F:growth factor activity; IDA.
 DR GO; GO:0005529; F:sugar binding; NAS.
 DR GO; GO:0008284; P:positive regulation of cell proliferation; IDA.
 DR InterPro; IPR001304; Lectin C.
 DR InterPro; IPR008994; Nucleic_acid_OB.
 DR Pfam; PF00059; Lectin C; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
 DR PROSITE; PS00641; C-TYPE_LLECTIN_2; 1.
 KW Direct protein sequencing; Glycoprotein; Growth factor; Lectin;
 KW Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 323 Stem cell growth factor.
 FT DOMAIN 183 320 C-type lectin.
 FT SITE 61 63 Cell attachment site (Potential).
 SQ SEQUENCE 323 AA; 35694 MW; D13604CDAF087427 CRC64;
 Query Match 23.0%; Score 220.5; DB 1; Length 323;
 Best Local Similarity 29.5%; Pred. No. 2.4e-11;
 Matches 57; Conservative 37; Mismatches 70; Indels 29; Gaps 7;
 QY 4 TQPKKIVNAKKDVVN-TKMFEEELKSRDLTLAQEVALLEKEQALQTVCLKGTKVHMKCFL 62
 DB 140 TQGLRQLNNAAGTDRDAVQALQEAQGRAE-----REHGRLEG-CLKGLRLGHKCF 189
 QY 63 APTQTKTFHEASEDCISRGGLSTPTQSGENDALYEYLRQSVG-NEAEIWLGLNDMAAEG 121
 DB 190 LSRDFEAAAQAARCTARGGSLAQPADRQMEALTRYLRAALAPYNWFWLGVHDDRAEG 249
 QY 122 TWDMVTGARIAYKNW----ETEITA-----OPDGKTKNCVLSGAANGKWF 165
 DB 250 LYLFENGORVSFFFAHRSRPRPELGAQPSASPHLSPDQPNNGTLENC-VAQASDDGSMWD 308
 QY 166 KRCRDQLPYICOF 178
 DB 309 HDCQRLYYVCEP 321
 RESULT 15
 PSPD BOVIN
 ID PSPD BOVIN STANDARD; PRT; 369 AA.
 AC P35246;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
 GN Name=SFTPD; Synonyms=SFTP4;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OC NCBI TaxID=9913;
 RN [1]_
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 208-247.
 RC TISSUE=Lung;
 RX MEDLINE=93170856; PubMed=8436402;
 RA Lim B.L., Lu J., Reid K.B.M.;
 RT "Structural similarity between bovine conglutinin and bovine lung
 RT surfactant protein D and demonstration of liver as a site of synthesis
 RT of conglutinin.";
 RL Immunology 78:159-165(1993).
 CC -!- FUNCTION: Contributes to the lung's defense against inhaled
 CC microorganisms. Binds strongly maltose residues and to a lesser
 CC extent other alpha-glucosyl moieties. It could participate in the
 CC extracellular reorganization or turnover of pulmonary surfactant.
 CC -!- SUBUNIT: Oligomeric complex of 4 set of homotrimers.
 CC -!- SUBCELLULAR LOCATION: Extracellular.

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2005, 11:53:10 ; Search time 39 Seconds
(without alignments)
446.544 Million cell updates/sec

Title: US-09-445-576A-7
Perfect score: 960
Sequence: 1 EPPTQPKKIVNAKDVNT.....KWFDRKRDQLPYICQFGIV 181
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues 283416
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	100.0	202	1 TTHUN	tetranectin precu
2	792	82.5	202	2 JC4031	tetranectin precu
3	408	42.5	166	1 A37289	tetranectin homo
4	207.5	21.6	369	2 S32603	surfactant protein
5	186.5	19.4	244	1 LNM5MC	mannose-binding le
6	186.5	19.4	1257	2 S28764	neurocan precursor
7	185.5	19.3	248	1 LNHUMC	mannose-binding le
8	182.5	19.0	1268	2 S52781	neurocan - mouse
9	180	18.8	883	2 S57653	brevican precursor
10	178.5	18.6	371	1 JN0450	conglutinin precu
11	178.5	18.6	371	2 I45878	conglutinin - bovi
12	177.5	18.5	244	1 LNR1MC	mannose-binding le
13	177	18.4	309	1 S34198	IgE FC receptor II
14	177	18.4	912	2 A54423	brevican precursor
15	176.5	18.4	374	1 A42046	surfactant protein
16	176	18.3	301	2 A53570	collectin-43 - bov
17	175	18.2	330	2 T46256	brevican - human (
18	172.5	18.0	375	1 A45225	pulmonary surfacta
19	172	17.9	883	2 A49126	brevican precursor
20	171.5	17.9	238	1 LNR1MA	mannose-binding le
21	166.5	17.3	239	1 LNM5MA	mannose-binding le
22	166	17.3	331	1 LNM5ER	IgE FC receptor, l
23	163	17.0	1456	1 A36563	mannose receptor p
24	160	16.7	1340	2 A39808	proteoglycan core
25	160	16.7	1479	2 T42710	mannose receptor,
26	160	16.7	2327	2 T42630	aggreacan - bovine
27	157.5	16.4	175	2 A37194	pancreatic thread
28	156.5	16.3	321	2 T26152	hypothetical prote
29	152	15.8	174	2 S54979	pancreatitis-assoc

30	151.5	15.8	321	2 T19259	hypothetical prote
31	151	15.7	1487	2 S48719	phospholipase-A(2)
32	151	15.7	2415	1 A39086	aggreacan precursor
33	150	15.6	2124	2 A28452	proteoglycan core
34	149.5	15.6	1326	2 B56395	secretory phosphol
35	149.5	15.6	1465	2 A56395	secretory phosphol
36	148	15.4	166	1 RGHU1B	regenerating islet
37	148	15.4	1643	2 T14274	versican precursor
38	148	15.4	2397	1 A55535	versican precursor
39	148	15.4	2409	1 A60979	versican precursor
40	148	15.4	3381	2 T42389	versican precursor
41	146.5	15.3	308	2 B88392	protein R06B10.3 [
42	146.5	15.3	321	2 T26153	hypothetical prote
43	146	15.2	2132	1 A55182	aggreacan precursor
44	145.5	15.2	1458	1 A49707	phospholipase A2 r
45	145.5	15.2	3562	2 A47171	chondroitin sulfat

ALIGNMENTS

RESULT 1

TTHUN

tetranectin precursor [validated] - human
N:Alternate names: plasminogen-kringle 4 binding protein
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1992 #sequence revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S24126; A56835; A29747; I38359; S19865
R:Berglund, L.; Petersen, T.E.

FEBS Lett. 309, 15-19, 1992

A>Title: The gene structure of tetranectin, a plasminogen binding protein.

A:Reference number: S24126; MUID:92380263; PMID:1511740

A:Accession: S24126

A:Molecule type: DNA

A:Residues: 1-202 <BER>

A:Cross-references: UNIPROT:P05452; EMBL:X70911

R:Wewer, U.M.; Albrechtsen, R.

Lab. Invest. 67, 253-262, 1992

A>Title: Tetranectin, a plasminogen kringle 4-binding protein. Cloning and gene expressi

A:Reference number: A56835; MUID:92365345; PMID:1354271

A:Accession: A56835

A:Molecule type: mRNA

A:Residues: 1-202 <WEW>

A:Cross-references: EMBL:X64559; NID:g37408; PIDN:CAA45860.1; PID:g37409

A:Experimental source: Placenta

R:Fuhlendorff, J.; Clemmensen, I.; Magnusson, S.

Biochemistry 26, 6757-6764, 1987

A>Title: Primary structure of tetranectin, a plasminogen kringle 4 binding plasma protein

A:Reference number: A29747; MUID:86107595; PMID:3427041

A:Accession: A29747

A:Molecule type: protein

A:Residues: 22-105, 'G', 107-202 <FUH>

A:Experimental source: Plasma

A>Note: S5-Ser and 58-Met were also found

C:Genetics:

A:Gene: GDB:TNA

A:Cross-references: GDB:135032; OMIM:187520

A:Map position: 3p22-3p21.3

A:Introns: 37/1, 70/1

C:Complex: homotetramer

C:Superfamily: tetranectin; C-type lectin homology

C:Keywords: glycoprotein; plasma; tetramer

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-202/Product: tetranectin #status experimental <MAT>

F:71-197/Domain: C-type lectin homology <LCH>

F:25/Binding site: carbohydrate (Thr) (covalent) #status experimental

F:71-81,98-197,173-189/Disulfide bonds: #status experimental

Query Match 100.0%; Score 960; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 4e-78;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 EPPTQPKKIVNAKDVVNTKMFELKSLDTLAQEVALLKEQALQTVCLKGTQVHMKC 81
QY 61 FLAFTQPKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db 82 FLAFTQPKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 180
Db 142 GTWVDMTGARIAYKNWETEITAPDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 201
QY 181 V 181
Db 202 V 202

RESULT 2

JC4031
tetranectin precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004
C:Accession: JC4031

R:Sorensen, C.B.; Berglund, L.; Petersen, T.E.

Gene 152, 243-245, 1995

A:Title: Cloning of a cDNA encoding murine tetranectin.

A:Reference number: JC4031; MUID:95137396; PMID:7835708

A:Accession: JC4031

A:Molecule type: mRNA

A:Residues: 1-202 <SOR>

A:Cross-references: UNIPROT:P43025; EMBL:X79199; NID:g671561; PIDN:CAA55791.1; PID:g6715

A:Experimental source: lung

C:Comment: This protein binds plasminogen, and may play a role in invasive cancer.

C:Superfamily: tetranectin; C-type lectin homology

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-202/Product: tetranectin #status predicted <WAT>

F:71-197/Domain: C-type lectin homology <LCH>

Query Match 82.5%; Score 792; DB 2; Length 202;
Best Local Similarity 80.7%; Pred. No. 3.6e-63;
Matches 146; Conservative 14; Mismatches 21; Indels 0; Gaps 0;

QY 1 EPPTQPKKIVNAKDVVNTKMFELKSLDTLAQEVALLKEQALQTVCLKGTQVHMKC 60
Db 22 ESTPTKAKAANAADLVSKMFELKRMNDVLAQEVALLKEQALQTVCLKGTQVHMKC 81
QY 61 FLAFTQPKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAE 120
Db 82 LLAFTQPKTFHEASEDCISQGGTLGTPOSELENEALFEYARHVSNDANIWLGLNDMAAE 141
QY 121 GTWVDMTGARIAYKNWETEITAPDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQFGI 180
Db 142 GAWVDMTGGLAYKNWETEITAPDGGKAENCAALSGAANGKWFDRKCRDQLPYICQFAI 201
QY 181 V 181
Db 202 V 202

RESULT 3

A37289
tetranectin homolog - reef shark
C:Species: Carcharhinus springeri (reef shark)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A37289; A37287
R:Neame, P.J.; Young, C.N.; Treep, J.T.
Protein Sci. 1, 161-168, 1992
A:Title: Primary structure of a protein isolated from reef shark Carcharhinus springeri
A:Reference number: A37289; MUID:93284081; PMID:1304877
A:Accession: A37289

A:Molecule type: protein

A:Residues: 1-166 <NE2>

A:Cross-references: UNIPROT:P26258

C:Superfamily: tetranectin; C-type lectin homology

C:Keywords: cartilage
F:37-160/Domain: C-type lectin homology <LCH>
F:37-47,64-160,136-152/Diulfide bonds: #status predicted

Query Match 42.5%; Score 408; DB 1; Length 166;
Best Local Similarity 44.5%; Pred. No. 4.3e-29;
Matches 77; Conservative 31; Mismatches 53; Indels 12; Gaps 2;

QY 6 KFKIVNAKDVVNTKMFELKSLDTLAQEVALLKEQALQTVCLKGTQVHMKCFLAFT 65
Db 2 KFSKSGKGKDD-----LRNEIDKLWREVNSLKENQALQTVCLKGTQVHMKCYLASR 52
QY 66 QTKTFHEASEDCISRGGTLSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAEAGTWWD 125
Db 53 GSKSYHAANEDCIAQGGTSLIPRSSDEGNSLSYAKSLVGARDFWIGVNDMTTEGKFDV 112
QY 126 MTGARIAYKNWETEITAPDGGKTENCACVLSGAANGKWFDRKCRDQLPYICQF 178
Db 113 VNGLPITYFNWD---RSKPVGGTRENCAVASTSGGKWSDDVCRSEKRYICEY 162

RESULT 4

S33603

surfactant protein D - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 17-Mar-1999

C:Accession: S33603

R:Lim, B.L.; Lu, J.; Reid, K.B.M.

Immunology 78, 159-165, 1993

A:Title: Structural similarity between bovine conglutinin and bovine lung surfactant prot

A:Reference number: S33603; MUID:93170856; PMID:8436402

A:Accession: S33603

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-369 <LIM>

C:Superfamily: pulmonary surfactant protein D; C-type lectin homology

F:248-367/Domain: C-type lectin homology <LCH>

Query Match 21.6%; Score 207.5; DB 2; Length 369;
Best Local Similarity 33.1%; Pred. No. 7.6e-11;
Matches 54; Conservative 27; Mismatches 61; Indels 21; Gaps 7;
QY 28 SRLDTLAQEVALLKEQ-QALQTVCLK-----GPKVHMKFLAFTQTKTPEASEDC 77
Db 216 AEVALQRVGLLEGQLQRLQNAFQYKAMLPNGRSVGEKIFKTVGSEKTFQDAQOIC 275
QY 78 ISRGTLSTPQTGSNDALYEYLRQSVGNEAEIWLGLNDMAAEAGTWWDTGARIAYKNWE 137
Db 276 TQAGGQLPSRSGAENEALQQL---ATAQNKAAPFLMSDTRKEGTFIYPTGEPIAVSNWA 332
QY 138 TEITAAP--DGGKTENCACVLSGAANGKWFDRKCRDQLPYICQF 178
Db 333 PQ---EPNNDGG-SENCVEI--FPNGKWNKDKVCGEQRLVICEF 369

RESULT 5

LNMSMC

mannose-binding lectin C precursor - mouse

N:Alternate names: Ra-reactive factor P28a

C:Species: Mus musculus (house mouse)

C:Date: 18-Jun-1993 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004

C:Accession: I48651; B46466; A42574

R:Sastri, R.; Wang, J.S.; Brown, D.C.; Ezekowitz, R.A.; Tauber, A.I.; Sastri, K.N.

Mamm. Genome 6, 103-110, 1995

A:Title: Characterization of murine mannose-binding protein genes Mb11 and Mb12 reveals f

A:Reference number: I48650; MUID:95284466; PMID:7766991

A:Accession: I48651

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-244 <SAS1>

A:Cross-references: UNIPROT:P41317; EMBL:U09016; NID:g773286; PIDN:AAA82010.1; PID:g77328

R:Sastri, K.; Zahedi, K.; Lelias, J.M.; Whitehead, A.S.; Ezekowitz, R.A.

J. Immunol. 147, 692-697, 1991

F/23-1257/Product: neurocan #status predicted <MAT>
F/176-253/Domain: link protein repeat homology <LNK1>
F/274-355/Domain: link protein repeat homology <LNK2>
F/364-366/Region: cell attachment (R-G-D) motif
F/953-984/Domain: EGF homology <EGF>
F/1029-1149/Domain: C-type lectin homology <LCH>
F/1156-1212/Domain: complement factor H repeat homology <FHD>
F/121,339,737,967,1164/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/372,410/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted
F/944/Binding site: chondroitin sulfate (Ser) (covalent) #status experimental

Query Match 19.4%; Score 186.5; DB 2; Length 1257;
Best Local Similarity 32.7%; Pred. No. 2.3e-08;
Matches 48; Conservative 20; Mismatches 52; Indels 27; Gaps 8;

QY 49 VCLAGT-----KVHMKCFIAFTQTTFHASEDCISRGGTLSTPQTSGENDALVEYL 100
DB 1021 LCEKDTGECDRGWKFGQHCYRYFAHRAWEDARDRCRRAGHLTSVHSPEEHKFI 1076
QY 101 RQSVGNENAEIWLGLNDMAAEG--TWVDMTGARIYKNWETETITAPQD---GGKTENCAV 154
DB 1077 NSFQHE-NSWIGLNDRTVERDFQWTDNTG--LQYENWREK---QPDNFFAGG--EDCVV 1127
QY 155 LSGAANKWFKDKCRDQLPYTCQFGIV 181
DB 1128 MVAHENGKRWNDVPCNYNLPYVCKGTV 1154

RESULT 7
LNHUMC
mannose-binding lectin precursor [validated] - human
N/Alternate names: mannan-binding protein
C/Species: Homo sapiens (man)
C/Date: 31-Mar-1989 #sequence revision 30-Sep-1991 #text change 09-Jul-2004
C/Accession: J01115; S05641; A34978; J00027; JX0319; PC2188; A32266
R/Sastry, K.; Herman, G.A.; Day, L.; Deignan, E.; Bruns, G.; Morton, C.C.; Ezekowitz, R.
J. Exp. Med. 170, 1175-1189, 1989
A/Title: The human mannose-binding protein gene. Exon structure reveals its evolutionary
A/Reference number: J01115; MUID:90010778; PMID:2477486
A/Accession: J01115
A/Molecule type: DNA
A/Residues: 1-248 <SAS>
A/Cross-references: UNIPROT:P11226; EMBL:X15422; NID:G34486; PIDN:CAA33462.1; PID:G34487
R/Taylor, M.E.; Brickell, P.M.; Craig, R.K.; Summerfield, J.A.
Biochem. J. 262, 763-771, 1989
A/Title: Structure and evolutionary origin of the gene encoding a human serum mannose-binding
A/Reference number: S05641; MUID:90073571; PMID:2590164
A/Accession: S05641
A/Molecule type: DNA
A/Residues: 1-248 <TAY>
A/Cross-references: EMBL:X15954; NID:G34480; PIDN:CAA34079.1; PID:G1212951
A/Accession: A34978
A/Molecule type: protein
A/Residues: 'X', 22-24, 'X', 26, 'X', 28-31, 'X', 33-34, 'X', 36, 'XXXX', 41-50 <TAY2>
R/Ezekowitz, R.A.B.; Day, L.E.; Herman, G.A.
J. Exp. Med. 167, 1034-1046, 1988
A/Title: A human mannose-binding protein is an acute-phase reactant that shares sequence
A/Reference number: J00027; MUID:86171281; PMID:2450948
A/Accession: J00027
A/Molecule type: mRNA
A/Residues: 1-2, 'C', 4, 'IT', 8, 'S', 10-57, 'R', 59-60, 'GT', 63-106, 'PGCLRR', 113, 'SSANRNGTYQ', 115
R/Kurata, H.; Sannoh, T.; Kozutsumi, Y.; Yokota, Y.; Kawasaki, T.
J. Biochem. 115, 1148-1154, 1994
A/Title: Structure and function of mannan-binding proteins isolated from human liver and
A/Reference number: JX0319; MUID:95073978; PMID:7982896
A/Accession: JX0319
A/Molecule type: protein
A/Residues: 1-248 <KUR1>
A/Accession: PC2188
A/Molecule type: mRNA
A/Residues: 1-20 <KUR2>
A/Experimental source: liver and serum
C/Comment: Mannose-binding lectins are opsonins that are important in host defense against

C;Comment: This protein is a Ca2+-requiring animal lectin specific for mannose and N-acetylglucosamine.
C;Gene: GDB:MBL
A;Cross-references: GDB:120167; OMIM:154545
A;Map position: 10q11.2-10q11.2
A;Introns: 63/1; 102/1; 125/1
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: acute phase; calcium binding; endoplasmic reticulum; Golgi apparatus; hydroxymethylglutaryl-CoA lyase
F;1-20/DNA: signal sequence #status predicted <SIG>
F;21-248/Product: mannose-binding lectin #status experimental <MAT>
F;42-98/Region: collagen-like
F;128-244/DNA: C-type lectin homology <LCH>
F;47,73,79,82,88/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

Query Match 19.3%; Score 185.5; DB 1; Length 248;
Best Local Similarity 30.2%; Pred. No. 5.4e-09;
Matches 45; Conservative 28; Mismatches 55; Indels 21; Gaps 5;

Qy 42 EQQALQTVCLK-----GTVHMKCFLAFTQTKTFHEASDCISRGCTLSTPTQSGE 92
Db 111 ERKALQTEMARIKKWLTFLSLGKQVGNKFFLTNGEIMTFKVKALCVKQASVATPRNAE 170
Qy 93 NDALYELRVSGVNEAEIWLGLNDMAAGTVDWMTGARIYKQWETETIAQP-DGKTEN 151
Db 171 NGALQNIKE-----EALGITDEKTEGQFVLDLTGNRLTYTNW---EGEPNAGSDSD 221
Qy 152 CAVLSGAANGKWFDPKRCRDLPYICQFIV 180
Db 222 CVLL-LKNGQWNVPCVSTSHLAVCFPI 248

RESULT 8
S52781
neurocan - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S52781
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, February 1995
A;Description: Amino acid sequence of mouse neurocan and brevican and their different exons.
A;Reference number: S52781
A;Accession: S52781
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1268 <RAU>
C;Cross-references: UNIPROT:P55066; EMBL:X84727; NID:g758629; PIDN:CAA59216.1; PID:g758629
C;Superfamily: aggregan; C-type lectin homology; complement factor H repeat homology; EGF
F;176-253/DNA: link protein repeat homology <LNK1>
F;274-355/DNA: link protein repeat homology <LNK2>
F;964-995/DNA: EGF homology <EGF>
F;1040-1160/DNA: C-type lectin homology <LCH>
F;1167-1223/DNA: complement factor H repeat homology <FHD>

Query Match 19.0%; Score 182.5; DB 2; Length 1268;
Best Local Similarity 31.8%; Pred. No. 5.4e-08;
Matches 47; Conservative 22; Mismatches 52; Indels 27; Gaps 8;

Qy 48 TVCLKGT-----KVHMKCFLAFTQTKTFHEASDCISRGCTLSTPTQSGENDALYEV 99
Db 1031 SLCEKDTGCDRGWHKFGQHCYRYFAHRAWEDAERDCRRAGHLTSVHSPPEHKFI--- 1087
Qy 100 LRQSVGNEAEIWLGLNDMAAG--TWVDMTGARIYKQWETETIAQPD-----GKTEACA 153
Db 1088 --NSFGHE-NSWIGLNDRTVERDFQWNTDNG--LQYENWREK---QPDNFFAGG--EDCV 1137
Qy 154 VLISGAANGKWFDPKRCRDLPYICQFIV 181
Db 1138 VVVAHESGRWVPCVNYLPPYCKGTV 1165

RESULT 9
S57653
brevican precursor - mouse

C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S57653
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, May 1995
A;Description: Cloning and sequence of mouse neurocan and brevican and their different exons.
A;Reference number: S57653
A;Accession: S57653
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-883 <RAU>
C;Cross-references: UNIPROT:O61361; EMBL:X87096; NID:g886889; PIDN:CAA60575.1; PID:g886889
C;Superfamily: aggregan; C-type lectin homology; complement factor H repeat homology; EGF
F;49-138/DNA: immunoglobulin homology <IMM>
F;173-250/DNA: link protein repeat homology <LNK1>
F;271-352/DNA: link protein repeat homology <LNK2>
F;626-657/DNA: EGF homology <EGF>
F;664-784/DNA: C-type lectin homology <LCH>
F;791-847/DNA: complement factor H repeat homology <FHD>

Query Match 18.8%; Score 180; DB 2; Length 883;
Best Local Similarity 28.0%; Pred. No. 5.9e-08;
Matches 45; Conservative 26; Mismatches 62; Indels 28; Gaps 5;

Qy 40 LKEQALQTVCLK-----TKVHM-----KCFLAFTQTKTFHEASDCISREG 82
Db 638 LEEKEGFRCLCLPGYGGDLCDVGLHFCSPGWEAFQGCYKHFSTRSWEESAQCRAUGA 697
Qy 83 TLSTPTQSGENDALYELRVSGVNEAEIWLGLNDMAAGTVDWMTGARIYKQWETETIA 142
Db 698 HLTSICTPEEQDFVNDYREYQ-----WIGLNDRTIEGDFLWSDGAPLLYNNW---PG 748

Qy 143 QPDGG--KTENCANLVSNAANGKWFDPKRCRDLPYICQFIV 181
Db 749 QPDSYFLSGENCVVMWHDQGSQSDVPCNVHLSYTCRMGLV 789

RESULT 10
JN0450
conglutinin precursor - bovine
N;Alternate names: C3b-binding protein
N;Contains: conglutinin-N
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
R;Suzuki, Y.; Yin, Y.P.; Makino, M.; Kurimura, T.; Wakamiya, N.
Biochem. Biophys. Res. Commun. 191, 335-342, 1993
A;Title: Cloning and sequencing of a cDNA coding for bovine conglutinin.
A;Reference number: JN0450; MUID:93213261; PMID:8460993
A;Accession: JN0450
A;Molecule type: mRNA
A;Residues: 1-371 <SUZ>
C;Cross-references: UNIPROT:P23805; DBJ:D14085; NID:g285643; PIDN:BAA03170.1; PID:g285643
A;Experimental source: liver
R;Kawasaki, N.; Itoh, N.; Kawasaki, T.
Biochem. Biophys. Res. Commun. 198, 597-604, 1994
A;Title: Gene organization and 5'-flanking region sequence of conglutinin: A C-type mammalian protein.
A;Reference number: JN0450; MUID:94128104; PMID:8297370
A;Accession: JN0450
A;Molecule type: mRNA
A;Residues: 1-371 <KA2>
A;Note: The authors translated the codon GAT for residues 250 and 270 as Glu
R;Lu, J.; Laurson, S.B.; Thiel, S.; Jensenius, J.C.; Reid, K.B.M.
Biochem. J. 292, 157-162, 1993
A;Title: The cDNA cloning of conglutinin and identification of liver as a primary site of
A;Reference number: S33235; MUID:93277452; PMID:7684896
A;Accession: S33235
A;Molecule type: mRNA
A;Residues: 1-172, 'H', 174-217, 'A', 219-271, 'V', 273-371 <LUJ>
C;Cross-references: EMBL:X71774; NID:g395267; PIDN:CAA50665.1; PID:g395268
A;Experimental source: liver
R;Lee, Y.M.; Leiby, K.R.; Allar, J.; Paris, K.; Lerch, B.; Okarna, T.B.
J. Biol. Chem. 266, 2715-2723, 1991

RESULT 11

I45878

conglutinin - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
C:Accession: I45878

R;Liou, L.S.; Sastry, R.; Hartshorn, K.L.; Lee, Y.M.; Okarma, T.B.; Tauber, A.I.; Sastry, Gene 141, 277-281, 1994

A;Title: Bovine conglutinin (BC) mRNA expressed in liver: cloning and characterization of A;Reference number: I45878; MUID:94215917; PMID:8163202

A;Accession: I45878

A;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A;Residues: 1-371 <L10>

A;Cross-references: UNIPROT:P23805; GB:L18871; NID:G495012; PIDN:AAA20126.1; PID:G495013

C;Superfamily: pulmonary surfactant protein D; C-type lectin homology

F;248-369/Domain: C-type lectin homology <LCH>

Query Match 18.6%; Score 178.5; DB 2; Length 371;
Best Local Similarity 31.7%; Pred. No.2.9e-08;
Matches 40; Conservative 23; Mismatches 58; Indels 5; Gaps 2;

QY 53 GTKVHKCFLAFTQTKTFHEASEDCISRGGLTSTPOTGSENDALYEVLRSQVGNAEIWL 112
Db 251 GQAVGEKIFKTAGAVKSYSDAEQLCREAKGQASPRSSAENAVTQVR---AQEKWYIL 307
QY 113 GLNDMAEGTWVDVTGARIAYKNWETEITAQPDGGKTENCANVLGGAANGKWFQKRCRDQL 172
Db 308 SMNDISTEGRTYPTGELTVYNWADGEPNNSDGQENCVEI--FPDGKNDVPCKSKQL 365

QY 173 PYICQF 178
Db 366 LVICEF 371

RESULT 12

LNRTMC

mannose-binding lectin C precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: A24791; A38322; JX0201; A26798

R;Drickamer, K.; Dordal, M.S.; Reynolds, L. J. Biol. Chem. 261, 6878-6887, 1986

A;Title: Mannose-binding proteins isolated from rat liver contain carbohydrate-recognition rotein.

A;Reference number: A24791; MUID:86196130; PMID:3009480

A;Accession: A24791

A:Molecule type: mRNA

A;Residues: 1-244 <DRI>

A;Cross-references: UNIPROT:P08661; GB:M14103

A;Note: part of the sequence, including the amino end of the mature protein, was confirmed by R;Childs, R.A.; Feizi, T.; Yuen, C.T.; Drickamer, K.; Quesenberry, M.S. J. Biol. Chem. 265, 20770-20777, 1990

A;Title: Differential recognition of core and terminal portions of oligosaccharide ligand A;Reference number: A38322; MUID:91065871; PMID:2249985

A;Accession: A38322

A:Molecule type: mRNA

A;Residues: 86,'EL',89-97 <CHI>

R;Wada, M.; Itoh, N.; Ohta, M.; Kawasaki, T. J. Biochem. 111, 66-73, 1992

A;Title: Characterization of rat liver mannan-binding protein gene.

A;Reference number: JX0201; MUID:92299655; PMID:1607365

A;Accession: JX0201

A:Molecule type: DNA

A;Residues: 1-244 <WAD>

A;Experimental source: liver

A;Note: the authors translated the codon CCA for residue 43 as Phe R;Oka, S.; Itoh, N.; Kawasaki, T.; Yamashina, I. J. Biochem. 101, 135-144, 1987

A;Title: Primary structure of rat liver mannan-binding protein deduced from its cDNA sequence A;Reference number: A26798; MUID:87194686; PMID:3032924

A;Accession: A26798

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-244 <ORA>

A;Cross-references: GB:X05023; NID:G56634; PIDN:CAA28687.1; PID:G56635

C;Comment: The molecule contains six identical chains, occurring as disulfide-bonded dimers. The molecule is a member of the lectin superfamily, and is specific for mannose.

C;Genetics:

A;Introns: 59/1; 98/1; 121/1

C;Superfamily: mannose-binding lectin; C-type lectin homology

C;Keywords: calcium; endoplasmic reticulum; Golgi apparatus; homohexameric; hydroxyproline

F;1-18/Domain: signal sequence #status predicted <SIG>

F;13-244/Product: mannose-binding lectin C #status experimental <MAT>

F;38-94/Region: collagen-like

F;124-240/Domain: C-type lectin homology <LCH>

F;29-34/Disulfide bonds: interchain #status predicted

F;69/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match

Best Local Similarity 18.5%; Score 177.5; DB 1; Length 244;

Matches 47; Conservative 24; Mismatches 61; Indels 15; Gaps 7;

36 EVALLK-EQALQTVCL--KGTVMKCFLAFTQTKTFHEASEDCISRGGLTSTPOTGSE 92

107 EIALRSELKAMRWLLSSENVGKKYFMSVRRMPLNRAKALCSELOQTVPATPRAEE 166

93 NDALYEYLRQSVGNEAEIWLGLNDMAAGTWDMTGARIYKNWETITAQPDG-KGTEN 151

167 NRAI-----QNAKDV-AFLGITDQRTENVFELDTGNRVYTNW---EGEPNVVSGEN 217

152 CAVLSGANGKWFDKRCRDLPYICOF 178

218 CVVL--LTNGKMDVPCSDSFLVVCBF 242

RESULT 13

S34198

IgE Fc receptor II, low-affinity - rat

N;Alternate names: CD23; lymphocyte IgE receptor

C;Species: Rattus norvegicus (Norway rat)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: S34198

R;Flores-Romo, L.; Shield, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayal

submitted to the EMBL Data Library, June 1993

A;Description: Inhibition of an in vivo antigen-specific IgE response by antibodies to C

A;Reference number: S34198

A;Accession: S34198

A;Molecule type: mRNA

A;Residues: 1-309 <FLO>

A;Cross-references: UNIPROT:Q63097; EMBL:X73579; NID:G313672; PIDN:CAA51981.1; PID:G3136

C;Superfamily: IgE receptor II; C-type lectin homology

C;Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr

F;1-25/Domain: intracellular #status predicted <INT>

F;14-22/Region: stop-transfer sequence

F;24-46/Domain: transmembrane #status predicted <TMM>

F;47-309/Domain: extracellular #status predicted <EXT>

F;126-309/Product: soluble IgE-binding factor (29K) #status predicted <IGI>

F;149-309/Product: soluble IgE-binding factor (25-27K) #status predicted <BFI>

F;164-283/Domain: C-type lectin homology <LCH>

F;192-283,260-274/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 18.4%; Score 177; DB 1; Length 309;

Matches 50; Conservative 43; Mismatches 71; Indels 44; Gaps 8;

4 TQPKKIVNAKDVNTK-MFEELSRDLTLAQEVALLEKQQA----- 45

87 SLSQNLNELQEDLNVKSNSELSQNLTLQEDLVNVKSNQGLNEKRAASLEKQEEV 146

46 ----LQTVCLKGTVM-----KCFLAFTQTKTFHEASEDCISRGGLTSTPOTGSE 91

147 AKLWIEILMSKGTACNVCPKDLHFQOKCYYPGEGSKQWIAKFTCSDLGRVLSHSQK 206

92 ENDALYEYLRQSVGNEAEIWLGLNDMAAGTWDMTGARIYKNWETITAQPDG-KGTEN 150

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-244 <ORA>

A;Cross-references: GB:X05023; NID:G56634; PIDN:CAA28687.1; PID:G56635

C;Comment: The molecule contains six identical chains, occurring as disulfide-bonded dimers. The molecule is a member of the lectin superfamily, and is specific for mannose.

C;Genetics:

A;Introns: 59/1; 98/1; 121/1

C;Superfamily: mannose-binding lectin; C-type lectin homology

C;Keywords: calcium; endoplasmic reticulum; Golgi apparatus; homohexameric; hydroxyproline

F;1-18/Domain: signal sequence #status predicted <SIG>

F;13-244/Product: mannose-binding lectin C #status experimental <MAT>

F;38-94/Region: collagen-like

F;124-240/Domain: C-type lectin homology <LCH>

F;29-34/Disulfide bonds: interchain #status predicted

F;69/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match

Best Local Similarity 18.5%; Score 177.5; DB 1; Length 244;

Matches 47; Conservative 24; Mismatches 61; Indels 15; Gaps 7;

36 EVALLK-EQALQTVCL--KGTVMKCFLAFTQTKTFHEASEDCISRGGLTSTPOTGSE 92

107 EIALRSELKAMRWLLSSENVGKKYFMSVRRMPLNRAKALCSELOQTVPATPRAEE 166

93 NDALYEYLRQSVGNEAEIWLGLNDMAAGTWDMTGARIYKNWETITAQPDG-KGTEN 151

167 NRAI-----QNAKDV-AFLGITDQRTENVFELDTGNRVYTNW---EGEPNVVSGEN 217

152 CAVLSGANGKWFDKRCRDLPYICOF 178

218 CVVL--LTNGKMDVPCSDSFLVVCBF 242

RESULT 13

S34198

IgE Fc receptor II, low-affinity - rat

N;Alternate names: CD23; lymphocyte IgE receptor

C;Species: Rattus norvegicus (Norway rat)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: S34198

R;Flores-Romo, L.; Shield, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayal

submitted to the EMBL Data Library, June 1993

A;Description: Inhibition of an in vivo antigen-specific IgE response by antibodies to C

A;Reference number: S34198

A;Accession: S34198

A;Molecule type: mRNA

A;Residues: 1-309 <FLO>

A;Cross-references: UNIPROT:Q63097; EMBL:X73579; NID:G313672; PIDN:CAA51981.1; PID:G3136

C;Superfamily: IgE receptor II; C-type lectin homology

C;Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr

F;1-25/Domain: intracellular #status predicted <INT>

F;14-22/Region: stop-transfer sequence

F;24-46/Domain: transmembrane #status predicted <TMM>

F;47-309/Domain: extracellular #status predicted <EXT>

F;126-309/Product: soluble IgE-binding factor (29K) #status predicted <IGI>

F;149-309/Product: soluble IgE-binding factor (25-27K) #status predicted <BFI>

F;164-283/Domain: C-type lectin homology <LCH>

F;192-283,260-274/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 18.4%; Score 177; DB 1; Length 309;

Matches 50; Conservative 43; Mismatches 71; Indels 44; Gaps 8;

4 TQPKKIVNAKDVNTK-MFEELSRDLTLAQEVALLEKQQA----- 45

87 SLSQNLNELQEDLNVKSNSELSQNLTLQEDLVNVKSNQGLNEKRAASLEKQEEV 146

46 ----LQTVCLKGTVM-----KCFLAFTQTKTFHEASEDCISRGGLTSTPOTGSE 91

147 AKLWIEILMSKGTACNVCPKDLHFQOKCYYPGEGSKQWIAKFTCSDLGRVLSHSQK 206

92 ENDALYEYLRQSVGNEAEIWLGLNDMAAGTWDMTGARIYKNWETITAQPDG-KGTEN 150

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-244 <ORA>

A;Cross-references: GB:X05023; NID:G56634; PIDN:CAA28687.1; PID:G56635

C;Comment: The molecule contains six identical chains, occurring as disulfide-bonded dimers. The molecule is a member of the lectin superfamily, and is specific for mannose.

C;Genetics:

A;Introns: 59/1; 98/1; 121/1

C;Superfamily: mannose-binding lectin; C-type lectin homology

C;Keywords: calcium; endoplasmic reticulum; Golgi apparatus; homohexameric; hydroxyproline

F;1-18/Domain: signal sequence #status predicted <SIG>

F;13-244/Product: mannose-binding lectin C #status experimental <MAT>

F;38-94/Region: collagen-like

F;124-240/Domain: C-type lectin homology <LCH>

F;29-34/Disulfide bonds: interchain #status predicted

F;69/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match

Best Local Similarity 18.5%; Score 177.5; DB 1; Length 244;

Matches 47; Conservative 24; Mismatches 61; Indels 15; Gaps 7;

36 EVALLK-EQALQTVCL--KGTVMKCFLAFTQTKTFHEASEDCISRGGLTSTPOTGSE 92

107 EIALRSELKAMRWLLSSENVGKKYFMSVRRMPLNRAKALCSELOQTVPATPRAEE 166

93 NDALYEYLRQSVGNEAEIWLGLNDMAAGTWDMTGARIYKNWETITAQPDG-KGTEN 151

167 NRAI-----QNAKDV-AFLGITDQRTENVFELDTGNRVYTNW---EGEPNVVSGEN 217

152 CAVLSGANGKWFDKRCRDLPYICOF 178

218 CVVL--LTNGKMDVPCSDSFLVVCBF 242

RESULT 13

S34198

IgE Fc receptor II, low-affinity - rat

N;Alternate names: CD23; lymphocyte IgE receptor

C;Species: Rattus norvegicus (Norway rat)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: S34198

R;Flores-Romo, L.; Shield, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayal

submitted to the EMBL Data Library, June 1993

A;Description: Inhibition of an in vivo antigen-specific IgE response by antibodies to C

A;Reference number: S34198

A;Accession: S34198

A;Molecule type: mRNA

A;Residues: 1-309 <FLO>

A;Cross-references: UNIPROT:Q63097; EMBL:X73579; NID:G313672; PIDN:CAA51981.1; PID:G3136

C;Superfamily: IgE receptor II; C-type lectin homology

C;Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr

F;1-25/Domain: intracellular #status predicted <INT>

F;14-22/Region: stop-transfer sequence

F;24-46/Domain: transmembrane #status predicted <TMM>

F;47-309/Domain: extracellular #status predicted <EXT>

F;126-309/Product: soluble IgE-binding factor (29K) #status predicted <IGI>

F;149-309/Product: soluble IgE-binding factor (25-27K) #status predicted <BFI>

F;164-283/Domain: C-type lectin homology <LCH>

F;192-283,260-274/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 18.4%; Score 177; DB 1; Length 309;

Matches 50; Conservative 43; Mismatches 71; Indels 44; Gaps 8;

4 TQPKKIVNAKDVNTK-MFEELSRDLTLAQEVALLEKQQA----- 45

87 SLSQNLNELQEDLNVKSNSELSQNLTLQEDLVNVKSNQGLNEKRAASLEKQEEV 146

46 ----LQTVCLKGTVM-----KCFLAFTQTKTFHEASEDCISRGGLTSTPOTGSE 91

147 AKLWIEILMSKGTACNVCPKDLHFQOKCYYPGEGSKQWIAKFTCSDLGRVLSHSQK 206

92 ENDALYEYLRQSVGNEAEIWLGLNDMAAGTWDMTGARIYKNWETITAQPDG-KGTEN 150

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-244 <ORA>

A;Cross-references: GB:X05023; NID:G56634; PIDN:CAA28687.1; PID:G56635

C;Comment: The molecule contains six identical chains, occurring as disulfide-bonded dimers. The molecule is a member of the lectin superfamily, and is specific for mannose.

C;Genetics:

A;Introns: 59/1; 98/1; 121/1

C;Superfamily: mannose-binding lectin; C-type lectin homology

C;Keywords: calcium; endoplasmic reticulum; Golgi apparatus; homohexameric; hydroxyproline

F;1-18/Domain: signal sequence #status predicted <SIG>

F;13-244/Product: mannose-binding lectin C #status experimental <MAT>

F;38-94/Region: collagen-like

F;124-240/Domain: C-type lectin homology <LCH>

F;29-34/Disulfide bonds: interchain #status predicted

F;69/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match

Best Local Similarity 18.5%; Score 177.5; DB 1; Length 244;

Matches 47; Conservative 24; Mismatches 61; Indels 15; Gaps 7;

36 EVALLK-EQALQTVCL--KGTVMKCFLAFTQTKTFHEASEDCISRGGLTSTPOTGSE 92

107 EIALRSELKAMRWLLSSENVGKKYFMSVRRMPLNRAKALCSELOQTVPATPRAEE 166

93 NDALYEYLRQSVGNEAEIWLGLNDMAAGTWDMTGARIYKNWETITAQPDG-KGTEN 151

167 NRAI-----QNAKDV-AFLGITDQRTENVFELDTGNRVYTNW---EGEPNVVSGEN 217

152 CAVLSGANGKWFDKRCRDLPYICOF 178

218 CVVL--LTNGKMDVPCSDSFLVVCBF 242

RESULT 13

S34198

IgE Fc receptor II, low-affinity - rat

N;Alternate names: CD23; lymphocyte IgE receptor

C;Species: Rattus norvegicus (Norway rat)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: S34198

R;Flores-Romo, L.; Shield, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayal

submitted to the EMBL Data Library, June 1993

A;Description: Inhibition of an in vivo antigen-specific IgE response by antibodies to C

A;Reference number: S34198

A;Accession: S34198

A;Molecule type: mRNA

A;Residues: 1-309 <FLO>

A;Cross-references: UNIPROT:Q63097; EMBL:X73579; NID:G313672; PIDN:CAA51981.1; PID:G3136

C;Superfamily: IgE receptor II; C-type lectin homology

C;Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr

F;1-25/Domain: intracellular #status predicted <INT>

F;14-22/Region: stop-transfer sequence

F;24-46/Domain: transmembrane #status predicted <TMM>

F;47-309/Domain: extracellular #status predicted <EXT>

F;126-309/Product: soluble IgE-binding factor (29K) #status predicted <IGI>

F;149-309/Product: soluble IgE-binding factor (25-27K) #status predicted <BFI>

F;164-283/Domain: C-type lectin homology <LCH>

F;192-283,260-274/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 18.4%; Score 177; DB 1; Length 309;

Matches 50; Conservative 43; Mismatches 71; Indels 44; Gaps 8;

4 TQPKKIVNAKDVNTK-MFEELSRDLTLAQEVALLEKQQA----- 45

Search completed: April 26, 2005, 11:53:57
Job time : 40 secs

THIS PAGE BLANK (USPTO)